Chapter 3
Risk Analysis in Investment Analysis: From Sensitivity Analysis to Time Series and Monte Carlo Simulation

Introduction – Risk Assessment

One of the tasks of any business is accounting for its historic financial results and keeping track of income and asset value in financial statements. Another much more challenging task faced by all but the very simplest of business endeavors is to make implicit or explicit forecasts of future cash flow that provide the basis for investment decisions. The problem with making forecasts is that they by definition encompass an unknown and uncertain future: with the exception of forecasting cash flow earned on risk free bonds, all cash flow forecasts will turn out to be wrong. The amount by which implicit or explicit forecasts of cash flow can be expected to be different from realized results forms the framework for thinking about risk in the pages below. This quite ambiguous concept of risk is not necessarily consistent with the attempt by most finance academics who suggest that risk can be captured in a single number such as the standard deviation or returns, beta of a stock, the volatility of prices or the value at risk of an investment portfolio. The question of whether risks can really be boiled down to such a single statistic or, alternatively, whether risk assessment is more of an art requiring a whole lot of judgment about business strategy, economics, and general experience is underneath many of the subjects in this chapter.

Financial modeling techniques discussed in the last chapter did not address how to quantify potential uncertainty in assumptions for key variables that drive projected revenues, operating expenses and capital expenditures. This chapter moves to the more difficult and interesting question of how to assess the uncertainty associated with future economic variables such as price, demand, expense structure, cost of new capacity and other items. After reviewing risk assessment techniques that rely on business savvy and economic judgment – sensitivity analysis, break-even analysis, scenario analysis and tornado diagrams – the remainder of this chapter describes how to create stochastic time series models statistics to measure risk. This discussion addresses how one can use a combination of mathematics and judgment in creating equations that allow one to model the future direction in prices (or other economic variables such as traffic volumes and interest rates) as well as the potential dispersion in those items. To develop time series models, a statistical tool kit of statistical parameters is introduced that includes volatility (the dispersion in prices), mean reversion (the speed at which prices come back to long-run average levels), correlations with other prices (such as the correlation between natural gas and oil prices), lower and upper price boundaries on price movements, price jumps, price trends and long-run equilibrium prices.

Evaluating the risk associated with future cash flows is at the heart of just about every issue in finance. The process of evaluating how much cash flow forecasts will differ from expected levels may involve estimation of how much estimated growth rates will be wrong; what will be the possible change in interest rates from expected levels; how much actual product prices and costs will be different from predicted prices and costs; what will be the change in public attitudes to various products; or what sudden changes in political events could occur. Attempting to measure risks such as these is by no means a new subject in finance and over the past fifty years Nobel Prizes have been awarded to

economists who have developed various "revolutionary" mathematical approaches to measuring risk ranging from beta to probability of default to value at risk. Notwithstanding the elegant formulas and complex statistics, application of new and seemingly promising innovative mathematical methods to risk assessment have generally turned out to be frustrating in practice and a conflicts continue to arise between whether risk should be evaluated using complex mathematical methods or business judgment. Given the amorphous nature of risk and basic questions about how to define risk, the difficulty of the quest to come up with mathematical methods to quantify risk is not surprising.

One of the things that the 2008-2009 financial crisis has highlighted is a general failure in the way financial professionals understand, quantify and model risk. Bankers were wrong in assessing the risk of sub-prime mortgages; equity investors misestimated the relative risk of stocks and bonds; rating agencies did not correctly measure the risk of collateralized debt obligations; and models developed by consultants to classify risks of loans so as to avoid failure of the financial system in the Basel II accord did not work at all. The difficulties in using complex techniques to measure risk are demonstrated in the failure mathematical models that quantify risk in a single number called value at risk. Value at risk is derived from sophisticated probabilistic analysis and it measures the amount of losses that will occur at something like a 99% confidence level over a given time period. The concept was used by banks and other companies to predict losses that could occur in collateralized debt obligations and many other investments. Problems with the value at risk statistic are described by David Einhorn, founder of a prominent hedge fund. He stated that value at risk statistics have been "relatively useless as a risk-management tool and potentially catastrophic when its use creates a false sense of security among senior managers and watchdogs. This is like an air bag that works all the time, except when you have a car accident."[1]

Alternative Ways to Assess Risk

These days, many corporations hire people with finance, mathematics, physics or economic degrees and give them a title of "risk manager." In one way or another, these people are supposed to evaluate how various decisions affect the risk and value of equity and debt investors. The somewhat mysterious job title seems to encompass everything from a junior analyst computing historic financial ratios for credit reviews to a person with a PhD in physics who develops value at risk statistics derived from complex mathematical equations. In practice, quantification of risk can encompass a range of analysis from simply graphing an economic variable (prices, costs, capital expenditures etc.) and observing how that variable affects cash flows and value -- sensitivity analysis -- to computing the probability distribution of cash flows using time series equations and Monte Carlo simulation. Various different types of risk analysis can be separated into different categories according to mathematical complexity. At one end of the risk assessment scale is a purely qualitative list and description of different risks. Other approaches moving along the mathematical complexity spectrum range from sensitivity analysis, to break-even analysis, to scenario analysis, to tornado diagrams, and finally to Monte Carlo simulation. The table below lists six different approaches to risk analysis along with a brief description of what is involved in each technique.

1 Nocera, Joe, "Risk Mismanagement" NYTimes http://www.nytimes.com/2009/01/04/magazine/04risk-t.html?pagewanted=print January 4, 2009.

Table 1

Risk Matrix	Sensitivity Analysis	Break-Even Analysis	Scenario Analysis	Tornado Diagram	Monte Carlo Simulation
List Economic and Financial Factors that can Influence the Financial Performance of an Investment and Determine whether the Risks are Mitigated through Contract Provisions and/or Hedging. For Variables that are not Mitigated, use Adjacent Risk Analysis Techniques	Choose an Important Economic Variable that is not Mitigated and make a Graph of how the Variable Effects the Outputs of Financial Variables Related to Valuation such as Net Present Value and IRR (The IRR can be the IRR on Equity, IRR on Debt or the Overall IRR of the Project Free Cash Flows)	Make a Table that Computes Outputs of Financial Variable Along Side of an Economic Variable and Determine when the Financial Variable Becomes Unacceptable. Once the Break-even Level is Established, one can Evaluate the Likelihood of the Economic Variable becoming the break-even Level.	Evaluate a Series of Different Output Variables Given a Set of Different Input Variables. For example, a Downside Case can be Established then Different Levels of Debt can be Assessed in the Downside Case. If the Downside Case cannot Support the Level of Debt, then Alternative Debt Levels are Recommened	The Tornado Diagram evaluates which variables have the most Significant Effect on the Financial Output of a Model Related to Valuation and which Variables have a Relatively Insignificant Impact. This tool can be used for Due Diligence Analysis or to Demonstrated the Relative Upside Potential and Downside Risk	Simulation Produces a Distribution of Returns and can Provide the Likelihood of Valuation Variables such as IRR and NPV Falling within a Certain Range. The Simulation Depends on Time Series Equations that are Derived from Parameters such as Volatility, Mean Reversion, Correlation and Price Boundaries

The above table puts judgmental approaches to risk assessment in the left columns and moves to the more mathematical methods on the columns listed at the right. A lot of the discussion in this chapter addresses Monte Carlo simulation -- the rightmost column of the table. The lack of old fashioned business judgment required in this approach does not at all imply that Monte Carlo models produce a better assessment of risk than careful business judgment regarding how a particular variable can move in the future. Very intelligent people who have made a large personal investment in the sturdy of complex mathematical approaches to risk assessment have a natural desire to apply their knowledge in practice. However most of the valuation mistakes discussed in Chapter one such as Eurotunnel and AES Drax would surely not have been solved through the more use of elaborate simulation and time series models. Indeed, some business valuation mistakes have been aggravated by inappropriately using historical data in applying complex models. The famous decision by Goldman Sachs in 2007 to limit its exposure to sub-prime mortgages demonstrates the potential efficacy of judgment over mathematics. The Goldman Sachs story of making decisions using judgmental opinion about how something "feels" rather than cold mathematical statistics is recounted by Joe Nocera:[2]

…Goldman called a meeting of about 15 people, including several risk managers and the senior people on the various trading desks. … They examined their VaR numbers and their other risk models. They talked about how the mortgage-backed securities market "felt." [According to Goldman Sachs' chief financial officer, David Viniar:] "Our guys said that it ***felt like*** it was going to get worse before it got better." The company made a decision to rein in the risk, which meant … getting rid of the mortgage-backed securities …. And that's why, back in the summer of 2007, Goldman Sachs avoided the pain that was being suffered by Bear Stearns, Merrill Lynch, Lehman Brothers and the rest of Wall Street.

A valuation nightmare that was not presented in the first chapter was the case of a company named Long-term Capital Management (LTCM). This case demonstrates problems with ignoring business judgment even more vividly than the Goldman Sachs story. LTCM was a hedge fund created in part by Nobel Laureates Myron Scholes and Bob Merton that, when it failed, almost brought the entire financial system to its knees in the late 1990's (although the problems seem like small change compared to

2 Nocera, Joe, "Risk Mismanagement" NYTimes http://www.nytimes.com/2009/01/04/magazine/04risk-t.html?pagewanted=print January 4, 2009.

Lehman Brothers and other problems of 2008.)[3] In addition to over-reliance on statistics, the LTCM case contained an explosive mixture of the valuation errors discussed in Chapter One that includes:

- Non-transparency in presenting financial information. Both investors and lending banks (the largest Wall Street Banks) were not allowed to see how trades were made and what models were used;
- Belief that staff employed by the fund was smart enough to continually beat the market and earn high economic returns in extremely competitive financial markets where risk adjusted returns should not be very high. The belief that the fund could continually earn high returns ultimately caused the fund to take higher and higher risks when returns began to fall;
- Investment strategies that were derived from “innovative” and highly complex mathematical models that supposedly discovered new ways to make money without being fully thought through. If LTCM could really make money by studying mathematical relationships (anathema to believers in efficient markets such as Scholes and Merton), then others would be able to copy the formulas and eliminate the profits;
- Faith in the reputation of other people -- Merton, Scholes and a famous bond trader named Bruce Meriwether -- without independently assessing whether the fundamental business concepts were viable.
- Investments that were made without verifying the underlying economics with simple back of the envelope checks. For example, the fund invested in Russian bonds by studying mathematical relationships rather than by visiting the country and seeing that military officers and secretaries were not being paid (in 1998 when Russia defaulted on its bonds, the country was arguably in a worse financial state than when the Soviet Union collapsed), and;
- Assumption that historic statistical relationships can be used to predict the future and be re-established after a period of time. LTCM failed when many economic variables did not act as they had in the past and sudden non-linear shocks occurred without historic precedent.

Notwithstanding all of these problems, the case raises a fundamental question of whether risk analysis can be reduced to a series of mathematical equations or whether risk must ultimately be assessed through old-fashioned business savvy. At best, LTCM demonstrated that complex quantitative techniques for risk measurement cannot be used without any supplemental business judgment. The healthy tension between relying on mathematical equations and using judgment raised in the LTCM case provides a backdrop in considering how best to assess different approaches to risk assessment.

Direct Risk Assessment versus use of Beta and WACC to Measure Risk

None of the methods shown in Table 1 above discussed the single parameter that finance theory implies can measure the relative risk of a stock, i.e. beta. Finance courses and textbooks suggest that one has to go no further than measure beta (which converts into a weighted cost of capital number) and then use this risk adjusted discount rate along with expected cash flows to measure value – no scenario analysis, no Monte Carlo simulation or any other direct measurement of risk is necessary. As beta supposedly contains all relevant information about risk, all of the techniques in Table 1 would not be necessary. Techniques discussed in this chapter are premised on the notion that risks cannot be magically be stuffed into one cost of capital number and involve a search for risk analysis methods other

3 Cite “When Genius Failed” and the Trillion Dollar Bet

than traditional discounted cash flow and weighted average cost of capital analysis. To contrast risk analysis using beta and weighted cost of capital with other approaches, consider the following two different hypothetical presentations made by different advisors to a company thinking about acquiring a company:

- The first advisor values the target company by carefully applying techniques taught in business school including computation of free cash flow, measurement of beta and use of on-going growth rates after the explicit forecast period. This method results in valuation of the target company on a standalone basis that is then adjusted for the value of synergies generated from the merger.

- The second advisor makes two financial models: one for the acquiring company without a merger and another for the combined acquiring company and target company using the integrated merger model techniques discussed in the last chapter. Then, after synergies are included in the analysis, he makes a presentation of whether earnings per share are higher in the combined case (accretion), or whether earnings per share decline (dilution) along with whether the company can maintain its investment grade bond rating. This integrated analysis accounts for the actual proposed financing of the transaction, the effect of the transaction on interest costs and stock prices as well as the fees paid for advisory services. The maximum value to be paid for the target company is the purchase price number that just avoids dilution in earnings per share. However in the presentation made by this second investment banker, there is no WACC to be computed, no terminal value to be estimated and the investment banker can apply similar assumptions for the acquiring company and the target company.

Knowing that valuations using the DCF technique can be so easily manipulated by making small changes in the WACC or the terminal growth rate – two variables which are very difficult if not impossible to measure – the company considering the merger most probably would pay more attention to the technique presented by the second advisor. Then, rather than pretending that risk of the making the acquisition can really all be incorporated in the WACC, the decision maker could ask for a series of different scenarios and break-even analyses that test how sensitive the accretion or dilution estimates are to various key variables with drive the forecasts. For example, he may evaluate the break-even point for the growth rate before which the merger becomes dilutive. Here, the financial officer is directly assessing risk rather than assuming it can be stuffed into the beta parameter. In a similar vein, decision makers assessing the risk of a project financed investment or a leveraged acquisition can directly assess risks to the equity IRR falling below a certain level and the coverage of debt service coverage using analogous techniques. These techniques to directly measure risk rather than adjusting the cost of capital are the subject of the remainder of this chapter (the subject of cost of capital and beta are covered in the next chapter.)

Judgmental (Non-Stochastic) Approaches to Risk Analysis

This section describes practical issues that arise in implementing the first five approaches to risk assessment listed in Table 1. After describing how the approaches work to analyze risk, a step by step process of how to implement the technique is presented. As with the discussion in Chapter two, you can skip the detailed instructions without missing the general idea of how the different methods work. All of the approaches use judgment relating to how high or low a variable can move without directly attempting to use probabilistic analysis.

Listing of Qualitative Risk issues and a Risk Allocation Matrix

A useful first step in directly assessing risk is to simply describe risks in some structured manner rather than diving into any quantitative analysis. There is no mathematical analysis whatsoever in this risk assessment technique. Instead, the process should force you to think about, discuss and understand a variety of different risks and how those risks could potentially affect investment returns. Qualitative analysis may involve discussing key risk issues, categorizing various risks into a matrix that describes risks or some other approach. Once the risks are defined and described, one can consider whether the some of the risks can be mitigated through one of three techniques -- insurance, contracts and or hedges. Mitigation of risks using one of these three methods generally involves some kind of explicit or implicit cost and an important valuation skill is the ability to consider tradeoffs between the benefits and costs of mitigating different risks (this type of analysis is addressed in Chapter Four.) Recall the Constellation case from Chapter 1, because of the lack of transparency in reporting, most analysts probably did not even identify the risk associated with declining market prices or the liquidity risk that arose after bonds were downgraded.

In most investments, some risks will not be mitigated by insurance, contracts or hedging. After all, if all risks could somehow be perfectly mitigated, then the investment would be risk free and values could be derived by simply discounting cash flows at the risk free interest rate. For those risks that are not mitigated, one apply one or more of the subsequent risk assessment techniques such as break-even analysis to evaluate the magnitude of the risk and whether the risk is acceptable for lenders and/or equity providers. Consider an extreme case where all of the risks are mitigated except for changes in the interest rate. Once the risk matrix demonstrates this fact, one determine how high interest rates would have to move before the investment cannot pay back its debt (from a lender perspective) or how high interest rates would have to rise before returns to equity holders fall below the risk free rate.

The table below illustrates the types of items that could be included in a risk matrix. The column on the left lists selected cases in which un-mitigated risks caused problems with the performance of investments.

Table 2

Risk Category	Description	Mitigation	Analysis of Un-Mitigated Risk	Example
Construction Phase				
Construction Over-run				
Material Prices		Contracts/Hedges	Break-even effect on IRR/Debt	Petro Chemical Plants
Cost Plus Provisions		Contracts	Cost of Comparable Projects	Eurotunnel/Eurodisney
Exchange Rates		Hedges	Break-even/Scenario/Monte Carlo	Petrozuata
Force Majure		Insurance		
Construction Delay				
Primary Project		Contracts-Liquidated Damage	Break-even effect on IRR/Debt	A380/US Nuclear Plants
Associated Projects		Contracts	Break-even effect on IRR/Debt	Eurotunnel
Technical Failure				
Resource Quantity		Resource Studies/Tests	Reserve Report	
Plant Efficiency		Contracts-Liquidated Damage	Engineering Report	Alstrom Combined Cycle
Operation Phase				
Price		Contracts/Hedges/Options	Break-even/Scenario/Monte Carlo	Argentina Merchant Plants
Volumes		Contracts	Break-even/Torndao Diagram	Uncongested Toll Roads
Cost		Contracts/Hedges	Break-even/Torndao Diagram	MCV Cogeneration Plant
Royalties		Contracts/Political Insurance	Break-even/Scenario/Monte Carlo	Petrozuata
Political				
Contract Abrogation		Political Insurance	Price Analysis/Off-taker Credit	Enron Dabhol/AES Drax
Nationalization		Political Insurance	Option Price Analysis	Petrozuata
Currency Convertibiltiy		Political Insurance		Zimbabuae Mines
Political Unstability		Political Insurance		Gaza Power Plant
Financial				
Interest Rate Changes		Hedges/Options	Break-even/Scenario/Monte Carlo	Pub Service New Hampshire
Inflation Rate Changes		Contracts	Break-even effect on IRR/Debt	PT Pation Indonesia
Re-financing Risk		Hedges/Options	Break-even/Scenario/Monte Carlo	Sub-prime Crisis

Sensitivity Analysis

One of the simplest and most effective ways to analyze risk is through simple sensitivity analysis. Sensitivity analysis simply means making an effective presentation as to what happens to some measure of value when the level of a key variable changes. High-powered analysts who compute statistics such as value at risk, probability of default and complex option price premiums, would probably scoff at sensitivity analysis as being overly simplistic. But consider a variable is very difficult to predict such as the level of traffic that will occur between England and France in the Eurotunnel, demand growth in the Philippines, the level of housing prices or electricity prices after a change in the structure of a market – all of which were discussed in the Chapter one valuation nightmares. Given the difficulty in creating an equation that can predict any of these variables, an effective picture of what happens to the value of the investment when the variable changes may be the best that one can realistically accomplish. For example, in assessing the value of collateral debt obligations that consisted of sub-prime mortgages, it surely would have been very useful to have a picture of what would happen to the value of the investments under alternative housing price assumptions. Alternatively, consider the case of Constellation Energy discussed in the Chapter one. Management presented value at risk statistics and maintained that the purpose of its trading activities was primarily to hedge other parts of the business (recall that management encouraged equity analysts simply to trend cash flows into the future.) This risk analysis frustrated investors and provided little useful information. With hindsight, a much more helpful presentation would have been a simple sensitivity analysis that illustrated what happens to Constellation cash flow and earnings under a wide range of energy price paths.

The figure below illustrates sensitivity analysis in the case of evaluating acquisition of an oil company, where acquisition uses senior debt, subordinated debt, preferred stock and common equity. A series of controls above the dashboard allows one to examine how projected cash flow is distributed using alternative financing structures and different operating assumptions. For example, through pushing the oil price up or down, one can observe the length time for the repayment of a loan, the rate of return earned on different securities and the value of the company and the point at which the a given rate of return cannot be achieved. For a senior debt, subordinated debt or equity investor, one could imagine decision makers varying the oil price and evaluating what happens to their cash flow. Rather than wasting a lot of time on forecasting oil prices, the investor could begin with his desired return and sees what level of oil price is required to attain the desired cash flow and the target rates of return. The picture is intended to demonstrate that sensitivity analysis is a lot about presentation and creatively summarizing key input drivers along with effective measures of value.

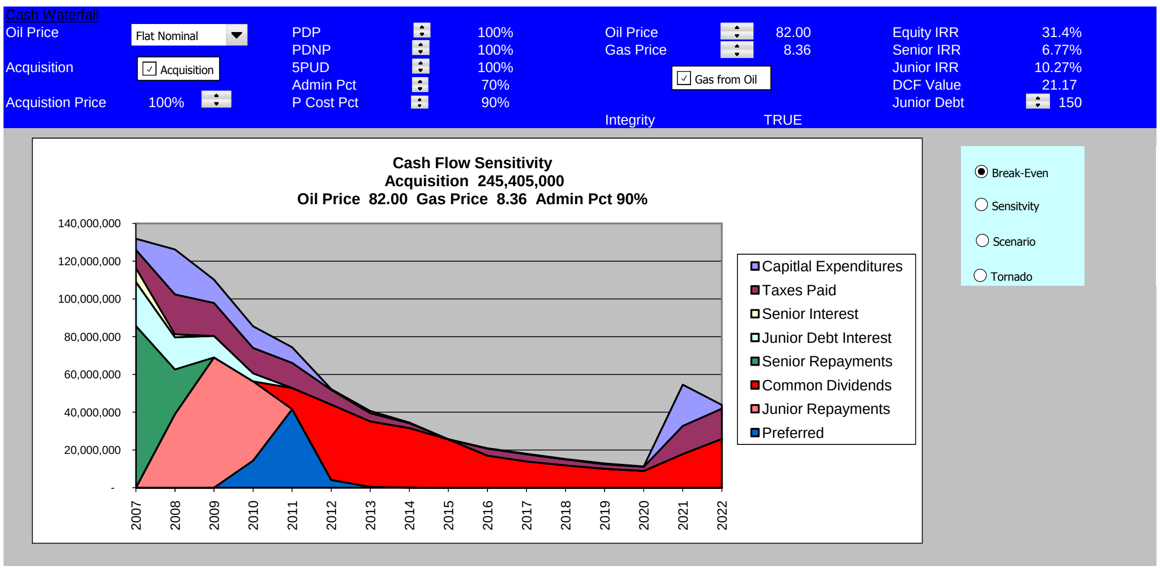


Figure 1

In order to quickly and effectively present important cash flow and valuation statistics along with economic and financial structuring variables, a few practical ideas are presented in the step by step analysis below.

Step 1: Select a variable from the financial model that presents cash flow or balances of a financial variable for presentation. Examples may include the balance of subordinated debt, earnings per share, cash flow to equity or cash flow and debt service.

Step 2: Collect the output variables in a separate section of the model (you could name this section of the model something like “graph data.”) If you structure the rows and columns in this part of the model with the name of the output variable in the cell that is immediately to the left of the data and the x-axis of the graph (e.g. the year) immediately above the data, then you can press the F11 key and very quickly make a graph of the data (the F11 key should quickly become one of your favorite short-cut keys.) To use the F11 key, simply select any single cell that is part of the data to be graphed and press the key. For the F11 key to work effectively, the bock of data should have blanks all around it – above the data,

below the data and to the left and right of the data. An example of setting-up data for graphing in this manner is shown below:

	2009	2010	2011	2012	2013
Total Revenues - Town	-	234,002	234,002	234,643	234,002
Total Operating Expenses	-	76,465	77,177	103,945	104,802
Total Capital Expenditure and WC	1,897,200	-	-	-	-
Total Debt Financing Costs	0	230,856	230,793	231,370	230,663
Total Debt Financing Benefits	1,897,200	-	-	-	-
Net Cash Flow to Town	-	-	-	-	-

If you would like to make a flexible graph in which the length of the graph or the selected data can vary, refer to sensitivity analysis portions of the excel background slides in the accompanying CD which describes how to create flexible named ranges using the OFFSET function in excel.

Step 3: The idea of presenting sensitivity analysis is to show how an input variable affects the output variables that are chosen for presentation with something like a spinner button. To do this, first choose one or more input variables that have an important effect on the rate of return, value or credit quality. These may be input variables such as the oil price in the example above, or they may be variables related to the structure of the transaction such as the purchase price or the amount of debt issued. To allow the variables to be effectively presented in the graph, you can use the spinner, scroll bar, check box or combo box forms available in excel. Unfortunately, there are a few quirks in excel that complicate this process, particularly when the graph is on a different page from the input data.

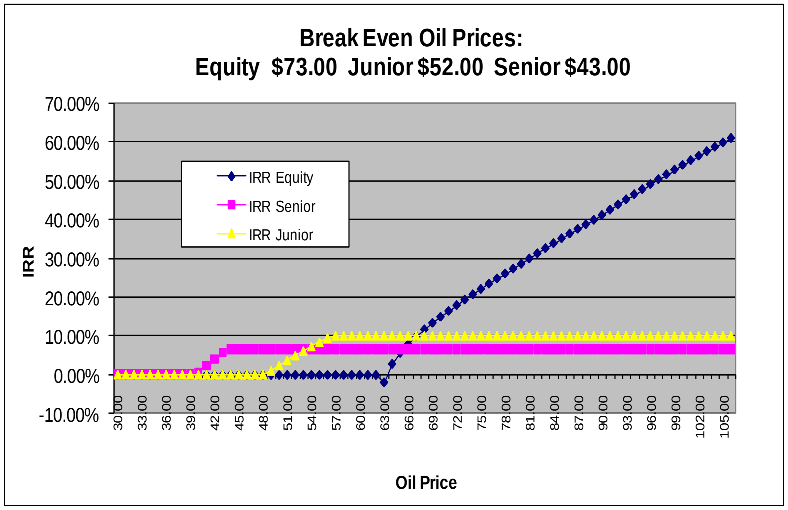
In order to create a form such as a spinner box, use the view, toolbars, forms menu in excel 2003, or you can use the insert controls tab from the developer menu in excel 2007 (make sure to use the control forms an not the activex controls) and insert the control on the sheet. Once the form is on the sheet, right click on the form and use the format controls option to attach the control to the input cell. However, when linking the cell, first click on another sheet and then click on the cell in the current sheet so that the sheet name is included in the cell link. Once you have created a form and attached it to an input cell in this manner, you can copy the form to other sheets and place the spinner box or other form directly in the graph. (If you are frustrated or even angry at this brief description, you can follow steps in the excel background files of the accompanying CD to work through the process more slowly.)

Step 4: It is often useful to present a title on the graph that contains both text and statistics that represent important valuation outputs such as the internal rate of return on the investment, the minimum debt service coverage ratio, or the net present value of cash flows. In addition, you may want to show some of the input variables that are being changed with the control forms (as shown in the graph above.) As with the above process above, there are a few quirks in excel that complicate the process. When combining text and numbers on a graph, you must first enter all of the information that will be included in the title into a single row (or column) somewhere in your file (you can change the format of the numbers and make the title as long as you would like.) Once the information has been entered in this manner, then select the title on the graph (there must be an existing title) and press the equal sign and refer to the single row or column where you entered the data. Next, press the enter key and the title should appear on the graph. If you would like the title on the graph to have multiple rows as in the example above, use the ALT and ENTER keys together at various cutoff points in the single row or column which contains the information. (As above, you can work through this slowly by referring the excel background files in the attached CD.)

Break-even Analysis

Break-even analysis is similar to sensitivity analysis, with a subtle but important difference that one can quantify risk into one single number which shows the cushion by which a variable can move before something bad or good happens. One of the supposed advantages of statistical measures of risk such as value at risk or the probability of default is that these numbers seemingly place all of the risk of an investment into a single number. By expressing risk in a single number, the break-even point does a similar thing, but in a more intuitive way. For example, consider the case of Eurotunnel described in Chapter one and pretend that you are an investor in one of the many tranches of junior debt. To evaluate the risk that your debt cannot be paid you could construct a financial model of the project and then keep pushing the traffic volume down until it reaches the level at which your debt cannot be repaid and you lose money. Once you have computed this break-even traffic level, if the chance that the actual traffic level will be below the break-even traffic level is very low, you could be confident that your debt will be repaid. On the other hand if the cushion between the base case traffic level forecasted by the consultant and the break-even level is thin, then the likelihood of default on your debt may be quite high. In evaluating alternative structuring possibilities such as different capital structures, purchase prices, covenant triggers, and other terms, it seems plausible that the investor would want to know how the break-even level of traffic if affected. Similar break-even measures could be developed for commodity prices, demand growth, cost structure and for different classes of investors – equity, senior debt, preferred stock and so forth.

As with sensitivity analysis, one of the advantages of using break-even analysis is that one does not have to make forecasts for variables that seem almost impossible to predict such as oil prices, growth rates, or traffic levels. Instead, the break-even level allows you to make a judgmental assessment as to whether the cushion between the current or expected level and the break-even level is wide enough to be comfortable with the investment. For example, if the break-even oil price is $20/barrel, one could simply compare this break-even level to historic real prices in assessing the potential for something bad to happen to the investment. Figure 2 below shows a presentation of break-even analysis using the oil transaction example above. The three lines on the graph show the rate of return on senior debt, subordinated debt and equity while the x-axis lists a range in potential oil prices. The point at which the variables fall below hurdle rates – in this case, the risk free interest rate, is the defined as the break-even point. If the current oil price is $85/barrel, then the equity break-even is quite thin, while the cushion for the senior debt is much higher. If the transaction used different levels of debt or equity, the break-even values would change. For example if more senior debt and less equity are issued to purchase the company, the break-even will be thinner for the senior debt.



Implementing break-even analysis is seemingly simple and it can indeed be though of as simply a subset of sensitivity analysis. For example, recall analysis subordinated debt in the oil acquisition case discussed in the context of sensitivity analysis. To determine the break-even oil price, one could simply push down the oil price until the IRR on the subordinated debt falls below the interest rate that was promised (the LIBOR rate plus the credit spread.) While general intuition behind a break-even analysis is simple, there are a couple of issues in computing and presenting sensitivity analysis that require some additional discussion. The first issue is establishing an effective criterion for purposes of deriving the break-even point. The second issue involves mechanical techniques to automatically present the break-even level whenever something changes in a financial model.

The process of establishing a criterion for determining the break-even point in a variable is often not obvious. Consider the perspective of senior lenders in a transaction where the most important risk variable is the projected growth rate. Here, the break-even analysis would evaluate the lowest acceptable growth rate which is acceptable for purposes of break-even analysis. One approach is to measure how low the growth rate can fall be before the minimum debt service coverage ratio becomes 1.0x – the level at which cash flow just covers debt service. The 1.0x criterion measures at what level cash flow is insufficient to meet the required payment, which seems to be a natural basis upon which to compute the break-even point. The minimum debt service coverage however may not represent the actual loss on debt as a company could be able to ultimately re-pay debt after missing a debt service payment. This means a better break-even analysis would be to see how low a variable can go until debt cannot be repaid at the end of a project. A third way to look at the issue is to evaluate how low the variable can fall until the IRR on debt falls below the risk free rate. This approach uses the notion that an investor has an option to either invest in a risk free security or risky investments and the break-even should shown what risks are taken to achieve a higher return than the risk free rate. Importantly, these three different criteria can lead to substantially different break-even points. Similar issues arise with respect to break-even points for equity investors, for acquirers in merger analysis, for government agencies and other decision makers. In the above graph, the break-even points are all determined by the point at which the IRR falls below the risk free rate.

Once the general procedure for deciding when the break-even point is established, the remaining issues are mechanical -- how to compute and present the break-even points. One method is simply to use the

goal seek tool in excel to derive the required level of the input variable that just meets the required output. For example, one could set the IRR on subordinated debt to the risk free rate by changing the estimated growth rate. A different approach is to compute tables that list increments of the break-even variable (e.g. the growth rate in demand) next to values produced of the criteria variables (e.g. the debt or equity IRR.) This method allows one to compute break-even points for different investments (and/or for different criteria) and present the break-even points without re-running the goal seek process multiple different times. The step by step instructions below explain how to compute break-even values using the second technique.

The first step is to create a one way data table where the sensitivity values are increments of the break-even variable is listed and a set of values for the criteria variables are computed for each of the increments. The data table tool in excel (accessed from the data menu in excel 2003 and from the what-if button in the data tab in excel 2007) is both one of the best and the worst features of excel. The good part is that you can quickly create scenarios and produce different break-even values for different transaction structures. The bad part is first that the data tables must be in the same sheet as the input variable – in this case the sheet that contains the oil price, the growth rate, the traffic demand or other break even variables. The second bad part is that data tables can slow down excel (unless the automatic except data table option is used.)

The data table used for the graph of oil prices and IRR's is shown on the figure below. The data table is created by setting up a restricted format where links to the criteria variable are placed one row above and one row to the right of the list of oil prices.

This is an input variable for the oil price that must be in this sheet

Set up the data table exactly like this with the formulas one row up and one to the right

	Input Variable - Oil Price			82
	Hurdle	6%	6%	6%
	Row Num	26	15	19
	Break Even	65.00	43.00	51.00
		IRR Equity	IRR Senior	IRR Junior
		31.4%	6.77%	10.27%
OIL PRICE SENSITIVITY	35.00	N/M	0.00%	0.00%
	37.00	N/M	0.00%	0.00%
	39.00	N/M	0.00%	0.00%
	41.00	N/M	2.40%	0.00%
	43.00	N/M	5.75%	0.00%
	45.00	N/M	6.77%	0.00%
	47.00	N/M	6.77%	0.00%
	49.00	N/M	6.77%	1.00%
	51.00	N/M	6.77%	3.76%
	53.00	N/M	6.77%	6.25%
	55.00	N/M	6.77%	8.58%
	57.00	N/M	6.77%	10.27%
	59.00	N/M	6.77%	10.27%
	61.00	N/M	6.77%	10.27%
	63.00	-1.91%	6.77%	10.27%
	65.00	5.69%	6.77%	10.27%
	67.00	9.95%	6.77%	10.27%
	69.00	13.44%	6.77%	10.27%
	71.00	16.51%	6.77%	10.27%
	73.00	19.44%	6.77%	10.27%

Once the data table is created, the remaining task is to find the break-even points for the alternative variables – in this case the equity IRR, the senior IRR and the subordinated or junior IRR. To find the break-even value for each criterion, two excel functions – the MATCH function and the INDEX function can be used together. The MATCH function finds the row number of the variable, in this case, the IRR, that corresponds to the criterion (in this case 6 %.) The INDEX function then can be used to find the break-even value that is associated with the row number. When the data table is created in this manner, different break-even values can be presented for different financial structures and different assumptions for variables that are not the break-even values.

While the break-even analysis presents risk as a single number, there are a couple of problems with the technique. First, the break-even analysis does not measure what happens when a cohesive set of variables that can move together. A rather obvious example is that in an extended recession, sales and prices may decline in tandem. All of the perfect storms that seem to occur much more than would be expected if variables would really move independently demonstrate this point. Second, the analysis depends on a single value of the break-even variable – for example, the oil price is assumed to stay at the constant level over the term of the modeling period. This means that use of a break-even analysis is difficult to accomplish when a variable is volatile and the price can have sharp upward and downward moves.

Scenario Analysis

Chapter one introduced the notion of debt capacity as a way to measure risk and ultimately derive the value of investments. In determining how much debt can be supported by an investment, there is no simple formula, because nobody really knows how banks and other lenders make decisions, but one oft-mentioned theory is that bankers come up with a downside case and then make sure there is some protection for them in this case. To illustrate this concept, consider the example of a wind farm in Germany where the projects receive a fixed “feed-in” tariff from the government and can hedge operations and maintenance risk using contracts. With energy prices and expenses hedged, the primary remaining risk is the possibility of wind variability. To determine how much debt a project such as this can support, a P90 case for wind speed is used – meaning that there is a 90% chance that the wind will be higher than this level. Then, to compute the amount of debt, the bank adds a 15% buffer on top of the expected cash flow in this downside case implying a DSCR of 1.15x. This example demonstrates how scenario analysis – in particular the downside case -- drives valuation decisions and real world risk assessment.

The difference between scenario analysis and the sensitivity and break-even techniques described above is that multiple different input variables are changed at the same time rather than focusing on a single variable. In developing a scenario analysis, one must come up with a comprehensive outlook with respect to a number of economic variables rather than simply seeing how much one variable can change before something bad happens. As discussed above, one of the most important scenarios is a downside case. If a company can easily survive a downside case – it can repay debt out of cash flow or easily re-finance debt maturities -- without cutting dividends and without changing capital expenditure policy, then the credit rating should be above investment grade. If the company cannot survive the downside case – perhaps meaning it cannot re-pay or re-finance maturing debt without dramatic changes in capital expenditures and other items -- then the credit rating should be below investment grade (BBB- in the Standard and Poor’s scale.)[4] In a project finance context, the downside case can be used to establish covenants, debt service reserves and other elements of a transaction as shown in the table below:

4 Fundamentals of Credit Analysis, McGraw Hill.

Scenario	Use in Project Fiance Model
Base Case	Establish the schedule for repayment of debt
Downside Case	Establish the level of debt service reserve Establish the covenant levels Develop repayment flexibility
Upside Case	Establish cash sweep mechanics Develop pre-payment structure

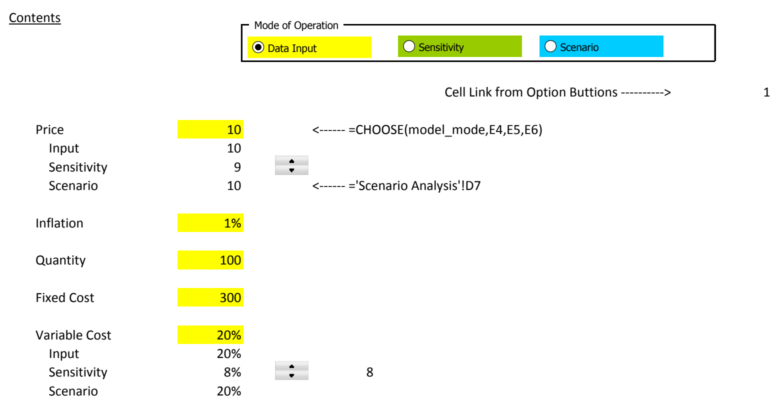
The obvious problem in using scenario analysis is determining assumptions for the level of various variables in the alternative cases. It is difficult enough to come up with base case assumptions which are intended to the expected or most likely outcome (not the budgeted case presented by company management which often contains an optimism bias related to completion of projects and implementation of business strategy.) In developing a downside case, one must not only come up with reasonable values for the variables, but also some sense of what the downside case is supposed to represent. For example, banks sometimes define the downside case as a 20% probability scenario, but nobody really has any idea of whether the set of variables really has a 20% chance of occurring. It is very easy to come up with a pessimistic set of variables; it is much more difficult to develop a set of consistent variables that are quite negative and have a relatively small chance of occurring. In the wind project example discussed above, the downside is relatively easy to develop because the risks of different wind conditions and technical functioning of a turbine can be observed from objective past data which should not deviate from historic data, meaning that future possibilities for wind speeds can be taken from a given distribution. On the other hand, construction of a downside case for the sub-prime mortgages would have been much more difficult to establish. To come up with this scenario one would have to project future housing prices, future income levels, the relation between housing foreclosure and income and many other economic variables. Unlike projection of the wind speed in the case of the German wind turbine, historical observations of past data may not be very useful in forecasting what could happen in the future. The difficulty in coming up with this sort of judgmental analysis required for scenario analysis is surely one motivation for developing mathematical representations of variables which supposedly measure the probability of the scenarios.

Modeling issues in developing scenario analysis involve structuring a model so that:

- Multiple different scenarios can be developed
- One can choose a host of different values
- One can make sensitivity analysis from different scenarios
- One can present compare different scenarios side by side in a table
-

Mechanical construction of scenario analysis poses fewer issues than coming up with the appropriate values to use in the analysis. One can use the scenario manager in excel, which allows you to change a number of different variables and create a report containing multiple output variables. Problems with this approach include: (1) new scenario pages must be created each time a change in the structure of the transaction is made and one must re-run the scenario manager each time the scenario manager is run; (2) the titles of the input variables and output variables must be manually adjusted; (3) the input variables must be entered into a data form rather than on a spreadsheet; (4) the scenario manager is difficult to manage when input variables are located on different pages; and (5) one cannot make drop down boxes to manage and graph the scenarios. Given these problems, an alternative way to develop scenarios using the INDEX function and the data tables is presented below on a step by step basis.

Step 1: Structure the assumptions part of the sheet to accept different sources of input variables, in this case from either the scenario analysis or as a direct input. One way to create a flexible model is to use the option buttons and conditional formatting as illustrated in the figure below. The option button produces a number which can be used together with the CHOOSE function to allow the model to operate in different modes (this variable is named “model_mode” in the figure below.) The grouping illustrated in the figure can be quickly created using the SHIFT, ALT, → short cut key (this is another very effective short-cut that is almost as good as the F11 key.) Note that this step is not essential if the model will always been run from a different set of scenarios.



Step 2: Enter data for different input variables in a scenario format. For example, enter the base case, the downside case, the upside case and a set of different sensitivity scenarios in a different rows and enter values for different assumptions (e.g. price, demand, cost structure) in separate columns. An example of entering scenario inputs is shown in the figure below that illustrates a completed scenario analysis (the example uses columns for the scenarios and rows for the input variables.) The input data is shown in the grey area in the figure below.

Step 3: Once the input data is entered for different scenarios, create a variable for the number of the scenario number and use INDEX function to find the input data associated with that scenario number. The use of the index function allows application of a data table because the whole scenario analysis is driven by a single number (and data tables are driven by a single number.) The scenario number is shown at the top of the table in the figure below.

Step 4: Model inputs should be driven from the different scenario data. This means that the cash flow model drivers should be linked to values computed from the INDEX function discussed above. The figure above illustrates this process as the scenario inputs come from another page that contains the scenario inputs and INDEX function.

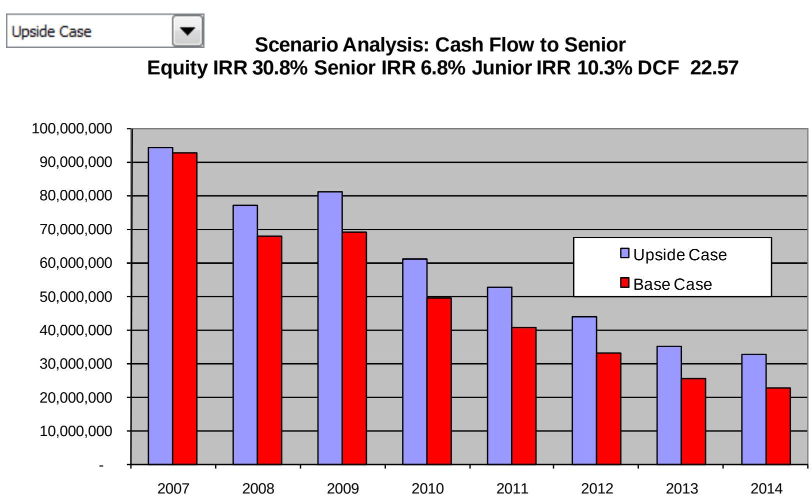
Step 5: Create a set of outputs alongside the scenario inputs. The outputs are structured by first entering a counter for the scenario number and then linking output variables one row above and one row to the right of the scenario numbers in order to allow construction of a data table. Once the output variables are structured in this manner, the data table tool can be used. An illustration of this process resulting in a scenario analysis is shown in the diagram below.

Transaction Cost	2,100.00
Scenario Number	3

Scenario Summary	Index Values	1	2	3
	Upside Case	Base Case	Downside Case	Upside Case
Price	12	10	9	12
Quantity	120	100	90	120
Inflation	3%	1%	-2%	3%
Fixed Cost	320	300	280	320
Variable Cost	19%	20%	22%	19%

Results			1	2	3
IRR		40.47%	17.91%	3.20%	40.47%
DCF		$2,687.62	$590.10	($421.92)	$2,687.62
Cash Flow		-2,100.00	-2,100.00	-2,100.00	-2,100.00
	1	846.40	500.00	351.80	846.40
	2	878.19	505.00	336.36	878.19
	3	911.00	510.05	321.15	911.00
	4	944.86	515.15	306.16	944.86
	5	979.80	520.30	291.38	979.80
	6	1,015.85	525.51	276.81	1,015.85
	7	1,053.06	530.76	262.45	1,053.06
	8	1,091.44	536.07	248.28	1,091.44
Payback		3.00	5.00	7.00	3.00

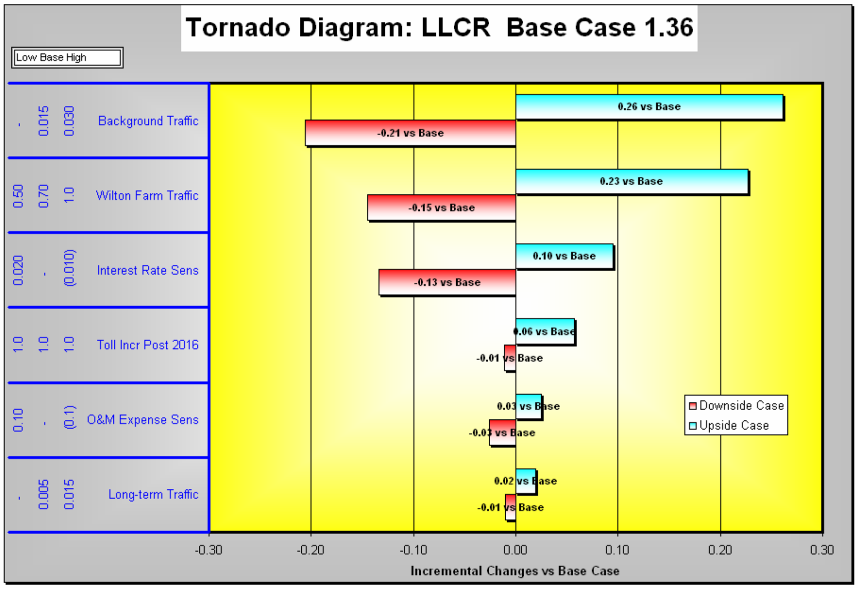
Once the scenario analysis is completed in this manner, various techniques can be used to present the scenario analysis. In the example below, the cash flow from the selected scenario is graphed along with the base case scenario and various financial data for the scenario is shown on the graph title. Techniques to develop this type of graph are described in the excel background folder included on the CD that accompanies this book.



Tornado Diagrams

The scenario analysis discussed above did not show which of the input variables are most important and which are not very important. This can be accomplished with a tornado diagram. A tornado diagram gets its name because it is supposed to look like the whirlwind created by a tornado which is larger on the top and smaller on the bottom. In a tornado diagram, input variables which have the highest effect on a selected output variable are shown at the top of the graph while variables which have a smaller impact are shown at the bottom. Through presenting sensitivity analysis in this manner, a tornado diagram can be used as a tool to determine which variables are most important in a sensitivity analysis and should be the focus of more study and due diligence. In addition, this type of presentation can be used to illustrate whether an investment has relatively more upside risk or downside exposure. Finally, a tornado diagram can be used as a tool to test the mechanics of a financial model. If the variables which have the highest effect are not intuitive, one should go back to the financial model and carefully examine the formulas related to the non-intuitive variable.

To interpret a tornado diagram, consider the figure below. Data is entered for a number of variables in base case, downside case and upside case scenarios just like the scenario analysis above. While the format is similar, the downside and upside cases do not necessarily have to represent a consistent outlook, meaning for example that if sales decrease, than operating costs may also decrease in a scenario analysis. When reviewing the effect of each variable, the downside case simply represents a low range outlook with respect to each variable. In the figure below, the data entered – on a judgmental basis -- for base case, downside case and upside cases are shown on the left hand side of the graph. Given these inputs, each bar on the diagram shows how the variation in the downside case from the base case, or the upside value relative to the base case affects the variable in question (shown on the top of the graph.) In the case at hand, the variable being studied is the loan life coverage ratio (“LLCR”) in a toll way project financing. The tornado diagram shows that given the input assumptions, the variable which has the highest effect on the LLCR is the background traffic. This variable causes the LLCR to fall by .21x, when all other assumptions use base case values, when the traffic growth is 0% rather than 1.5% (the downside case and base case assumptions shown on the left of the graph.) Similarly, if the background traffic increases by 3% per year, the LLCR increases by .26x relative to the base case. On the other hand, variables such as the operation and maintenance expense and the long-term traffic projections have a relatively small effect on the LLCR, meaning that if the LLCR is an important measure for analysis, one would not have to worry a great deal about these variables as long as the judgmental downside and upside assumptions are reasonable.



Mechanical implementation of a tornado diagram involves development of a two way data table where each variable is changed one at a time and the scenarios are also changed. When isolating the effect of one variable, all of the other variables are held at the base case value. In the above example, the background traffic is changed from the base case to the low case to the high case while all of the other variables are held at the base case value. When the base case for any variable is computed, the output is always re-set to the base case value as all of the other variables also remain at the base case number. In order to create a tornado diagram, you can use a file included on the attached CD that contains a number of macros. Using this file, you can add a tornado diagram into any financial model. Instructions for implementing the model are included in the file. Alternatively, and probably more useful, you can create your own tornado diagram using the following process:

1. Set up a scenario analysis using the INDEX function together with the data table as described above, where the column is used in the INDEX together with the scenario number. This has two objectives. First, it allows you to maintain the scenario analysis. Second it provides a basis for creating data for the tornado diagram.
2. Add a form control that allows one to select between a scenario analysis and a tornado diagram (You can use option buttons where you insert a button and then select a cell link. The option buttons work well with the CHOOSE function in excel).
3. Add an input for a variable number and enter numbers across each of the variables. The variable number can be entered near the scenario number that was used in the scenario analysis and the number for each variable can be entered above or below the variable title.
4. Below the INDEX computation, add a variable that includes a true/false test for whether the variable in the table is the same as the variable number input. For example, say there are six variables as illustrated in the diagram above. There should be a list of numbers from one to six above the variables. The test evaluates whether the input number for the variable number equals the number associated with the variable. There is only one value that is TRUE for the list of variables. The formula for the test should use the F4 key for the variable number and compare this to each formula as illustrated below:

   Variable Number Input (Fixed with F4) = Number Associated with Each Variable

5. Create an IF function using the test variable. If the variable is true, then use the value from the INDEX function discussed in step 3 above. If the variable is false, then use the base case value. Using this approach, every variable remains at the base case value except the variable which equals the variable number input. The IF test has the form illustrated below:

   IF(TRUE/FALSE Test, INDEX from Step 3, Base Case from Scenario Table)

6. Use the CHOOSE variable together with the cell link described in step 2 to select either the INDEX value – for the scenario option, or the IF function – for the tornado option. The values in this part of the analysis should drive the financial model. An illustration of this part of setting up a tornado diagram is illustrated in the excerpt below. The last line in the table is used in the input section of the financial model.

Scenario Number	2
Variable Number	2
Shares Issued	5,000.00
Premium	40%

Scenario 2 | Tornado | Run Table

	Trafic per Store	Margin per Store	Cap Exp per New Store	Growth Rate in Stores	Cost Synergy Percent		IRR on Equity - PE Ratio	IRR on Equity - M/B Ratio	Overall IRR on Free Cash Flow	5-Year Average Accretion	DCF Value - Equity
							12.78%	13.49%	9.36%	10.12%	17,769.60
	1	2	3	4	5						
Base Case	230,000.0	30.0%	850,000	4.50%	10.00%	1	13.3%	14.0%	9.8%	12.4%	19,049.43
Low Case	210,000.0	29.0%	930,000	3.00%	0.00%	2	5.8%	7.3%	5.1%	-13.0%	7,479.95
High Case	250,000.0	33.0%	800,000	6.00%	15.00%	3	18.3%	18.7%	14.1%	29.4%	32,683.48
Worst Case	210,000.0	28.0%	930,000	2.00%	0.00%	4	3.9%	6.0%	4.4%	-14.3%	6,728.93
Index Function	210,000.0	29.0%	930,000	3.00%	0.00%	=INDEX(H12:H15,Scenario_Number)					
Test	FALSE	TRUE	FALSE	FALSE	FALSE	=$D$4=H11					
If Test	230,000.0	29.0%	850,000	4.50%	10.00%	=IF(H18,H17,H12)					
Low Case	230,000.0	29.0%	850,000	4.50%	10.00%	=CHOOSE($I$4,H17,H19)					

7. After setting up the variable using both IF test that runs one variable at a time, set up a two way data table where one of the variables represents the variable number and a second variable represents the scenario number for the base case, low case and high case. The form of this two way data table is illustrated below:

		Variable Number				
		1	2	3	4	5
Sc Num	1					
	2					
	3					

8. In the above table, the first row represents the base case, the second is the low case and the third is the high case. In terms of columns, each column contains the value of the selected output variable if the input variable is changed and every other variable remains at the base case value. Once the table is set-up, run the data table either using the excel data table function or a macro as described in the appendix at the end of this chapter.
9. For the first row, the base case output is simply repeated for in each case. Once the table is computed, the base case versus the low case and the base case versus the high case can be computed by simply subtracting the low case row from the base case row and by subtracting the high case row from the base case row. This presents the impacts of different variables which is the basis for the tornado diagram. All that is left to do is to present the data. An illustration of this step is shown below. You can see that the variable with the largest effect if

variable number 5 and the second largest is variable number 2 with variable number 3 coming in last place.

		Variable Number				
		1	2	3	4	5
Sc Num	1	13.3%	13.3%	13.3%	13.3%	13.3%
	2	11.4%	12.8%	12.9%	12.8%	10.8%
	3	14.9%	14.8%	13.6%	13.9%	14.5%
Low vs Base		-1.9%	-0.6%	-0.5%	-0.5%	-2.6%
High vs Base		1.6%	1.5%	0.3%	0.5%	1.2%
Absolute Value		3.5%	2.1%	0.8%	1.0%	3.7%

10. In presenting the data, the idea of a tornado diagram is to sort the absolute value of the variables in the above table from the smallest to the largest. However, use of the sort function in excel is not helpful in this instance for a variety of reasons. First you probably do not want to change the order of all the variables in the scenario table. Second and more importantly, if you have used the data table function, you cannot sort items in the data table.
11. To use a more flexible sorting function, the first step is to create a single variable that will be a sort key used for tabulating the variables in a sorted order. Once the sort key is established (with the MATCH function), you can use the INDEX function to present selected variables in sorted order.
12. To sort a variable one can use the SMALL function or the LARGE function. In this case use the SMALL function with the absolute value shown in the table above. The mechanics of the small function is illustrated below:

SMALL(Fixed Row to Sort (Absolute Value from Table Above), Variable Number)

13. Match the sorted variable against the original absolute value line so as to establish a sort key that can be used in with the INDEX function to extract sorted variables. When computing the sort key, use the MATCH function so as to create an exact match. To do this enter a zero as the match type as illustrated below:

MATCH(Sorted Single Value, Unsorted Values -- Fixed, 0)

14. To create a graph, you will need the title of each variable as the x-axis and the low versus the base and the high versus the base. To set-up the data for making a graph, use the INDEX function along with the sort key. For example, to find the title, enter the original titles (fixed with the F4 key) in the index command and use the sort key as the column number (you do not need a row number) as illustrated below:

INDEX(Original Series of Titles – Fixed, Sort Key)

	Cap Exp per New Store	Growth Rate in Stores	Margin per Store	Traffic per Store	Cost Synergy Percent	=INDEX($D$10:$H$10,H37)
Low vs Base	-0.49%	-0.51%	-0.57%	-1.92%	-2.56%	=INDEX($D$32:$H$32,H37)
High vs Base	0.27%	0.51%	1.50%	1.57%	1.16%	=INDEX($D$33:$H$33,H37)

15. Graph the sorted variables using the F11 function and change the chart type to a stacked bar chart (make sure that there is no caption next to the titles so the excel will know that this is the x-axis.

Stochastic Risk Analysis and Time Series Equations

The remainder of this chapter discusses application of mathematics and statistics to risk assessment. A benefit of applying statistical analysis to risk assessment is that all of the above judgmental approaches depended in one way or another on the opinions of people as to the assumption of how high or how low some variable may become in the future. These opinions can be biased upwards by managers having a favorable attitude towards an investment concept or biased downwards if people generally disagreeing with a business idea. Such human biases can theoretically be avoided by using cold hard statistics to come up with the range in values for the upside and downside case variables. In fancier language, the process of using statistical analysis rather than judgment to measuring risk can be termed applying a stochastic diffusion processes to prediction of economic variables. One way to think about this stochastic analysis is that ranges in variables for sensitivity analysis, break-even analysis and scenario analysis come from objective statistical analysis of historic data rather than from subjective judgment.

While there are benefits associated with stochastic analysis in terms of providing objectivity to risk analysis, there are certainly also some disadvantages. One problem is the fact that history is often a very poor predictor of future economic variables that are driven in part by human behavior. Economic variables such as price and demand growth are distinguished from physical variables such as wind speed, reserves of oil in a field or hydro conditions that do not depend on the whims of human beings. A second problem is that parameters required to implement models including volatility, mean reversion, correlation, price trends, boundary levels, and jumps are often almost impossible to measure from historic data. Thirdly, there can be volatility in long-term equilibrium prices, lower or upper boundaries or changes in volatility itself can make the whole process boil down to purely random numbers. Fourth, general assumption that rates of changes in variables follow a normal distribution is often not valid which renders the whole process biased if not useless. Finally, even with modern software and fast computers, the process of implementing Monte Carlo simulation can be time consuming.

Given all of these problems, you may wonder why so much of this chapter is devoted to the subject of stochastic modeling. The general idea behind describing the mathematical approach to risk analysis in some detail is not necessarily to advocate the approach or to suggest that you should go out and immediately apply time series and Monte Carlo simulation. Instead, reasons for becoming familiar stochastic risk analysis techniques include: (1) making sure that you will not be intimidated or overly impressed when presentations are made using the approaches; (2) encouraging you to fully understand the flaws as well as the benefits of the stochastic modelling techniques; (3) explaining how the mathematical techniques can be combined with business judgment to make the process applicable in real world situations; and (4) using the stochastic representations of economic variables as a framework to think about how key variables can potentially move in the future.

The fourth point above is worthy of a bit more elaboration. To implement a time series model for a variable, one must come up statistical parameters that represent long-term trends, variations around the trend, eventual reversion to long-run production costs, lower and upper boundary values, possible sudden moves, and the relationship between the variable and other variables. Computing these parameters, and more importantly, thinking carefully about them, forces one to think about how key variables can potentially move in the future. For instance, in establishing downside case and upside case assumptions one should think about what is the lower limit, how much can the variable move in a year, will the variable move back to a long-term equilibrium level, what is that equilibrium level, can a

"perfect storm" cause a dramatic change in the variable, and how does the variable move with other variables. Thinking about economics is also important from the stochastic modeling perspective. When implementing parameters in time series equations, you need to understand the underlying economic factors which drive parameters such as volatility, mean reversion parameters so that you can better understand the real sources of risk in an investment.

In attempting to convey the advantages and disadvantages of stochastic analysis in risk assessment, the remainder of this chapter is organized into five sections. After describing some general terminology associated with stochastic risk analysis, a simple example of process is presented. This simple example demonstrates how it is very easy to construct a time series equation and apply Monte Carlo simulation to compute statistics such as the probability of default, value at risk and distributions of returns. (You can complete a Monte Carlo simulation in excel in a matter of minutes.) After presenting this simple case, the discussion moves to subtle but important issues in constructing alternative types of time series equations that incorporate alternative volatility, mean reversion, correlation, price boundaries and price jumps in economic variables. The fourth section describes how to use historic data in measuring parameters for a time series model – volatility, mean reversion, price boundaries, correlations and jump processes. The final section describes how to apply time series models in making price forecasts with a Monte Carlo simulation.

Terminology and Judgment in Stochastic Modeling

As with many subjects in finance, defining a few of terms makes stochastic modeling of risk easier to understand and explain. A few of the terms that define the stochastic modelling process are described below before working through more detailed analysis. These include time series equations, volatility, normal distributions, mean reversion and Monte Carlo simulation. Before delving into details of how to implement and analyze stochastic models, a discussion of how to combine the process with judgment.

Definition of Time Series Equations, Volatility, Mean Reversion and Monte Carlo Simulation

Time series models are mathematical formulas that describe an economic variable in terms of its dispersion as well as the expected level of the variable. For example, the formula for price in a financial model described in the last chapter may be $P_t = P_{t-1} \times 1.1$. This simply means that the current year's price is last year's price increasing at a rate of 10%. An equation such as this can be termed a deterministic equation. If the equation was a time series model, then the next year price would be modeled with a possible dispersion – say the price may increase by 20%, by 10%, by 0% or by percentages in the range. The expected value is still a 10% increase, but the time series equation provides probabilities of achieving different values within the range. The manner in which the price can vary around the expected value of an increase of 10% is defined by a volatility parameter. If the volatility is close to zero, there is little chance of the price moving by much more or much less than 10%. If the volatility parameter is high, the price can move in a wide range.

Through using time series equations rather deterministic equations in a model allows one to project ranges in values rather than only focusing only on one case at a time. By contrast, stochastic modelling incorporates the potential variance of cash flows into the model through inclusion of a volatility parameter. This allows all of the outputs of the model such as the IRR, the DSCR, the amount of debt that is repaid and any other variable to be expressed as a probability distribution. Among other things, time series equations with volatility of cash flow drives the probability of default in credit analysis; it

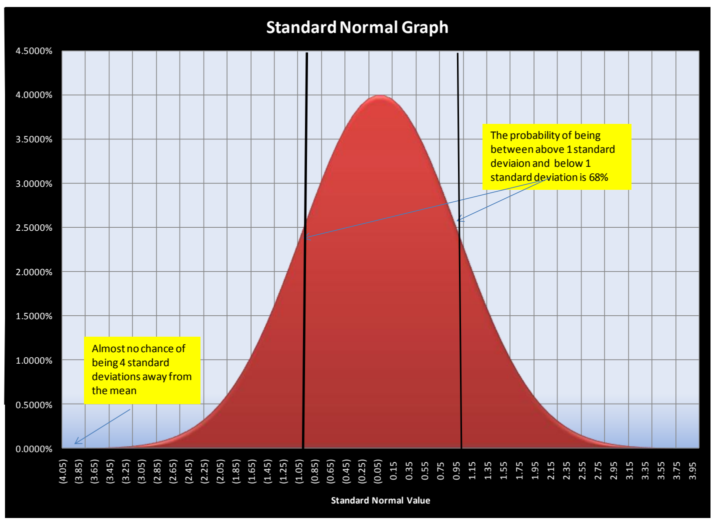
determines the value of real options; it is the reason for hedging risks through long-term contracts; and it is a key variable in determining value at risk. In common parlance, volatility is expressed as a percentage. In rough terms, volatility can be defined as the standard deviation of the rate of return. The rate of return can be the rate of change in stock prices, oil prices, demand, cost or any other data.

If the percent change in a variable comes from a normal distribution, volatility can be used to measure the probability that future values will be above or below a certain level. To see this, it is necessary to remember a little bit of statistics that most of us forget. In a normal distribution, the probability of being within one standard deviation of the mean is 68.27% and the probability of being within two standard deviations is 95.4%. For example, if the mean is 4 and the standard deviation is 3, then there is a 68.27% chance that the observed value will be between 7 and 1 – i.e. within one standard deviation of the mean above and below. The normal distribution is so convenient to use because any probability can be obtained if one knows only the mean and the standard deviation. Since volatility is the standard deviation of change in a variable, if one knows the volatility, the probably of achieving a value for the next year can be estimated. Say the volatility of oil prices is 20% per year and the average oil price in January is $54 per barrel. Then there is a 68% chance that the oil price will be between $43.2 and $64.8 (an increase or decrease of 20%) by the end of the year. Similarly, there is a 95% chance that the oil price will be between $75.6 and $32.4. These are the approximate actual values for 2007 oil prices and volatility. By the way, the price at the end of 2007 was $96 per barrel, well outside of the 95% range.

From a mechanical standpoint, it is worthwhile understanding the mechanics of a couple of excel functions that compute the probabilities of normal distributions. In working with normal distributions, there are two functions that are useful in determining whether the rate of change in variables really comes from a normal distribution. The first function uses the mean, the standard deviation and an observation from the distribution as inputs and then returns the probability of being less than the observation. For example, say the mean rate of change is 2%, the standard deviation is 15% and you would like to know the probability of realizing a rate of change of less than 20%. To find this number you could use the excel NORMDIST functions with arguments of 20%, 2%, 15% and a switch of 1 which signifies use of a cumulative distribution (NORMDIST(20%,2%,15%,1).) In this case the probability of achieving a growth rate of 20% or less is 88.5%. If a distance of one standard deviation above the mean of 17% is used, then the probability is 84.13%, while if the observation is one standard deviation below the mean -- -13%, then the probability is 15.87%. The difference between these values found with the NORMDIST function (84.13% minus 15.87%) yields the 68% chance of being within one standard deviation of the mean.

A second function that is useful when working through mechanical issues associated with stochastic modelling is a function in which you input the probability and you can find the number in the distribution associated with this probability. Using the numbers from the previous example, suppose we would like to find the percent change that has a 95% or a 99% probability, meaning that there is a 95% or a 99% change that the actual value will be below the computed value. This is the essential concept of value at risk. In the above case with a 2% mean and 15% distribution, one can use the NORMINV function. The arguments for this function are the probability, the mean and the standard deviation. To find the percent change for which we can be sure that 99% of the time we will be below the number, one would enter NORMINV(99%,2%,15%). The resulting percent change is 37%. Using the oil price example, if the volatility is 20% and we want to be 99% sure that the price will be below a computed value, we would find the percent change in price is 47%, implying a 2007 year end price of $79.19 per barrel. Recall that the actual price was $96. With a beginning price of the year price of $54, a volatility of 20% and an ending price of $96, one can use the NORMINV to show that the probability is 99.996%.

Since any value of a normal distribution can be expressed as the mean and standard deviation, one can subtract the mean from each value of a distribution and divide the result by the standard deviation. This number gives you the distance from the mean for any value. This number is known as the standard normal distribution. In the earlier example, where the mean was 2% and the standard deviation was 15%, the standard normal value of 20% is (20%-2%)/15% or 1.2, meaning that the 20% value is 1.2 standard deviations from the mean. The probability of this value can be found with the NORMSDIST function (the S is included for the standard normal distribution), where one simply enters the standard normal value. A graph of the standard normal distribution created with excel is shown below (as usual, press the F11 key). To create this graph, simply enter the standard normal values beginning with -4 and increasing to +4 (the chance of a normal distribution falling outside of this range is tiny.) Once the numbers are entered in a column, use the NORMDIST distribution with a mean of zero and a standard deviation of one and a switch of zero so the graph is not cumulative. This graph will be used below to evaluate whether distributions approximate a normal distribution.



The above definition of a time series equation discussed how the dispersion in a variable can be modeled using a volatility parameter. In addition to volatility, time series equations may include parameters such as mean reversion, trend factors, boundaries, jumps, long-run equilibrium values and correlations with other prices to project how variables evolve over time. Mean reversion refers to the tendency of a variable, after it increases above the mean, to move back to the mean. For example if the long-term mean of the oil price is $65 per barrel, and the current price is $100 per barrel, then the price may have a tendency to move back to the mean value. The speed at which a variable moves back toward the mean is called the mean reversion factor which can be added to a time series equation. Other analogous factors can be developed for trend parameters, correlation parameters and so forth. Details of how these parameters are included in time series equations and alternative methods for computing the parameters is described in subsequent sections.

Once a time series equation is defined and a probability distribution such as the normal distribution is selected, a financial model can be used to create probability distributions of model outputs. Monte Carlo simulation was originally used in development of the Atomic Bomb in the 1940's and it gets its name from the notion random numbers that are created by playing a roulette wheel many times at

Monaco.[5] The general idea of Monte Carlo simulation is to construct a large number of possible scenarios by using multiple random numbers together with a time series model. The random numbers are given structure because each random number is converted to a probability (using something a process like the NORMINV function) and then the probability is applied in the time series equation.

Combining Judgment and Historic Statistics in Stochastic Modeling

Statistical concepts including time series equations and Monte Carlo simulation discussed in this chapter have been applied for decades in valuation of financial securities including equities, bonds, foreign exchange and other financial instruments. The famous Black-Scholes model that is so often used in all sorts of valuation applications assumes asset prices follow a stochastic process and come from a normal distribution. After the Black-Scholes formula became very widely accepted, the subject of applying mathematical analysis to real non-financial investments as well as financial investments gained popularity. There is now a lively debate as to whether stochastic analysis can be realistically applied in cash flow modeling for valuation. By working through various examples in this chapter, you can hopefully assess for yourself the efficacy of this technique.

Part of the attraction of using stochastic time series equations is to remove judgment from the process as described above. Some who have studied hard in university and who work with stochastic models and computer programs such as “at risk” develop arrogance about non-stochastic models and look down on analysis that does not incorporate mathematical analysis as intellectually inferior. They come up with statistics such as probability of loss, earnings at risk and produce many pictures of the distribution of financial ratios that would not be possible with judgmental analysis. This attitude is naïve and dangerous as removing judgment from predicting how economic variables such as returns on stocks, commodity prices, growth rates, demand levels and other variables will move very often simply cannot be applied understanding the underlying economic and business factors that drive the variables. The controversy in over-relying on models and not using judgment is explained in the following statement by Alan Greenspan, a person who once had a lot of credibility and now may not be taken quite so seriously: “The essential problem is that our models – both risk models and econometric models – as complex as they have become – are still too simple to capture the full array of governing variables that drive global economic reality. A model, of necessity, is an abstraction from the full detail of the real world.”

One alternative to this is to combine judgment and with statistical analysis in creating time series equations. This sort of approach is suggested by a project finance analyst named Paul Ashley: “Given plausible assumptions for key risk drivers, the [credit] ratings are then built around simulating cash flows and debt coverage for individual projects. These simulations can then determine probability and loss distribution and directly observe the impact on cash flows and riskiness of a particular transaction structure under different market environments. The parameterization of the rating simulations combines market data and expert judgment in a structured way that allows transparency and consistency in risk evaluation without losing the ability to capture the specifics of a transaction structure.”

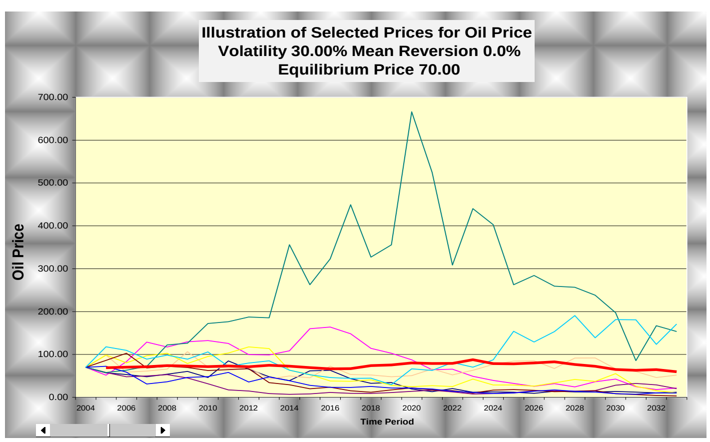
As with many statements made by financial professionals, the above statement requires a little translation. When the author uses the term simulation is used in the above quote, he means that some kind of time series equation is used together with Monte Carlo simulation. The word parameterization in the excerpt means converting predictions of key variables in a financial model into parameters such as volatility and mean reversion. The notion of transaction structure implies that once a Monte Carlo

[5] See Jackel, Peter, “Monte Carlo Methods in Finance”, 2001, for a technical description of Monte Carlo simulation.

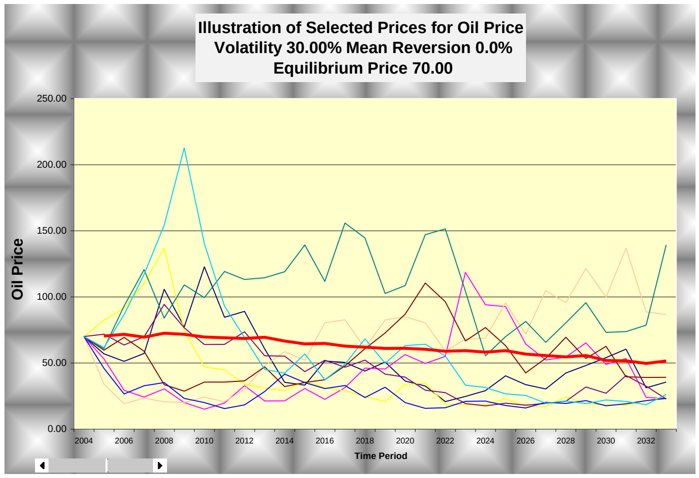
simulation has been created resulting in thousands of scenarios, this scenario data can be run with alternative debt levels, covenants, purchase prices and other factors to see how the simulation affects the probability of defaulting on debt (somewhat like running alternative transaction structures with scenario analysis discussed above.) Finally, the comment that expert judgment can be used means that it is possible to adjust time series parameters and to add factors such as lower and upper boundaries into an analysis which are derived from judgmental thinking about how low or high variables can move in the future.

The manner in which judgment can be incorporated into stochastic analysis is somewhat analogous to the break-even discussion above. Recall that break-even analysis involves coming up with the minimum or maximum level of a variable that allows the transaction to work. For example, one would come up with an oil price and then compare that price to historical trends, other forecasts, long-term productions costs and other factors. Similarly, the numbers produced by simulation should be inspected to see if they are consistent with historical observations of the variable and your judgment. A few specific examples of including judgment in a time series analysis include:

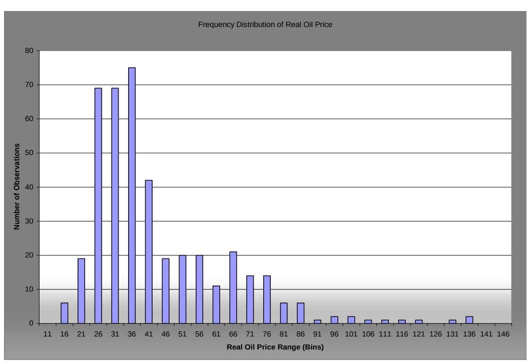
- Look at a number of a number of the scenarios for price, demand or other variables that are modeled with Monte Carlo simulation and evaluate the ranges from a judgmental perspective. For example, if many of the scenarios result in a price of zero or an unrealistically high price, the volatility, mean reversion and other parameters should be revised. The graph below shows selected scenarios that result from an assumption of no mean reversion, no upper or lower bounds and 30% volatility. Many of the scenarios result in a price of zero and some result in a price of more than 600. Clearly some judgment must be used to adjust the prices.

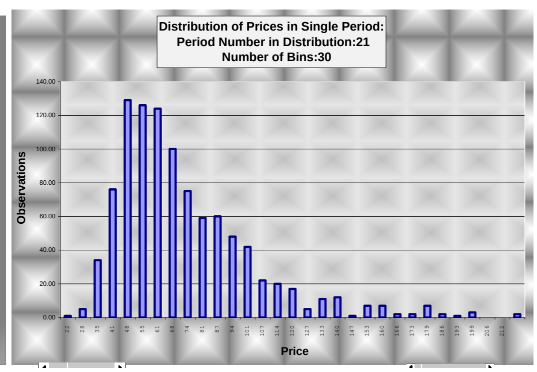


- Impose upper and lower boundaries on a judgmental basis. In the above example, one might estimate the short-run marginal cost to be 20 and the amount before which people stop consuming at 200. This judgmental adjustment of the parameters causes a dramatic difference in the simulations as shown in the graph below.



- Examine the frequency distribution of the variables that results from the Monte Carlo simulation and compare the simulated distribution with the actual historic distribution. If the simulated distribution has a much different appearance in terms of skewness and variation than the historical distribution, then you should go back and further study the time series parameters. The example below compares a simulated distribution with mean reversion to an actual distribution of oil prices. Simulated prices are shown on the chart on the right and the distribution of actual prices is shown on the left.





The mechanics of creating a frequency distribution shown in the above graph involve using the FREQUENCY function in excel. This function is called an array function because the output goes into multiple cells instead of a single cell. To accomplish this, first list the ranges in the variable that will be shown on the x-axis of the graph. These x-axis ranges are named bins. Once numbers for the bins are entered, shade the area next to the bins and then, while the range is shaded, type the function =FREQUENCY(data,bins). Arguments for the frequency function are first the data being evaluated and then the bins into which the data will be summarized. For example, if there are 70 observations of the oil

price between 35 and 40, then the result of the frequency function will be 75. This function will be used again below in evaluating whether numbers come from a normal distribution.

- Later in the chapter we will discuss including a jump process in time series equations. Adding this to a model may sound sophisticated, but it simply means that there is some probability (usually pretty small), that something really good or really bad will happen. The suggestion that these very unlikely events can be measured from statistical analysis and/or analysis of historic data is simply disingenuous. The whole idea of a sudden an very unexpected jump (like the effect of the Lehman Brothers collapse on the S&P 500) is that the event is unexpected and that it has not happened before. It is hard to imagine how a jump process can be added to time series models without some business judgment.

- Some of the judgmental techniques described above can be combined with stochastic modelling where one performs sensitivity, break-even or scenario analysis on parameters such as volatility and correlation to examine how sensitive risk analysis is to alternative parameters in time series models.

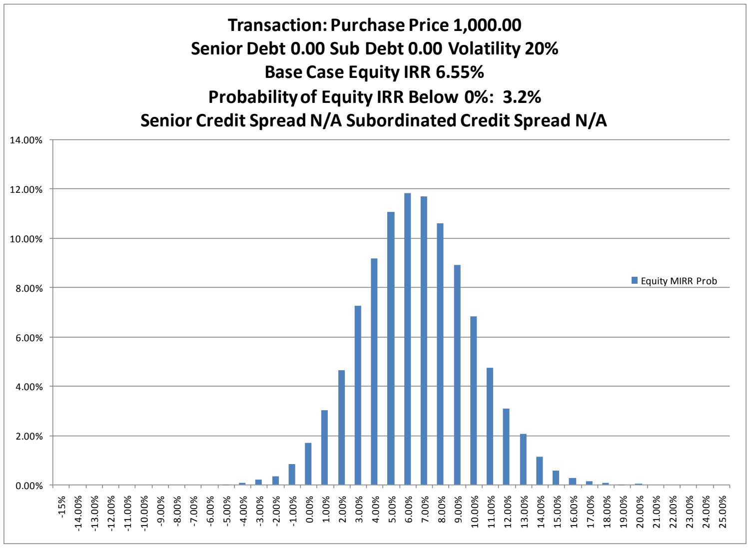
Overview of Time Series, Monte Carlo and Risk Measurement using a Simple Example

Working through a simple example of how one can easily create risk measures with Monte Carlo simulation rather than jumping into the complexities of time series models should remove the mystery of the method. Needless to say, the fact that the mechanics of applying Monte Carlo simulation are so easy does not mean that you should add Monte Carlo simulation to all of your financial models. The idea of the exercise is to assure that you are not intimidated by the mechanics and the general idea of Monte Carlo simulation but also to warn you of the dangers simplistically applying statistical concepts to risk analysis. To understand what is involved with Monte Carlo simulation, it may be easier to make a simulation than to read the concepts.

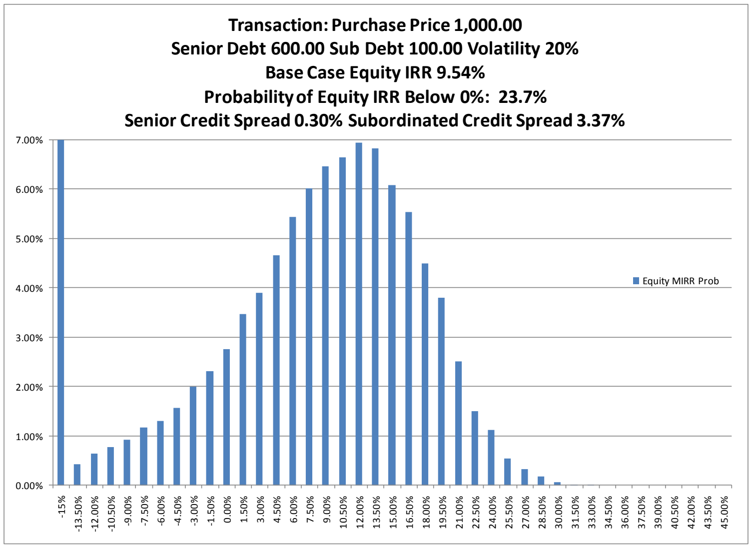
The step by step discussion below shows how within a few minutes you can integrate seemingly sophisticated risk statistics into financial models. The graphs at the bottom of this paragraph show results of an analysis that show the probability of earning more than the risk free rate and the required credit spread on alternative types of debt using different volatility parameters and different transaction structures. To do this analysis, you do not need fancy computer software or a PHD in econometrics. Instead, the only requirements are a basic understanding of volatility, knowledge of how to construct a simple time series model and a blank excel sheet. The analysis uses a boring investment of 1,000 with operating cash flow is 150 that has a ten year life. Without financing and without volatility the modified IRR is 6.55% assuming a risk-free re-investment rate of 5%. So far, we know nothing about the whether this is a good investment because we have no information about its risk.

To add risk to the investment, assume the cash flows have a volatility of 20%. In this case assume that the investment is financed completely with equity. By including only this factor, converting the 150 of cash flow into a time series equation and using Monte Carlo simulation, one can now answer the questions like what is the probability that the equity IRR will be below zero. The graph below shows the distribution of IRR's that result when the volatility of cash flows is included in the analysis. The graph demonstrates that it is possible to earn as much as 16% and that the probability of earning a return of

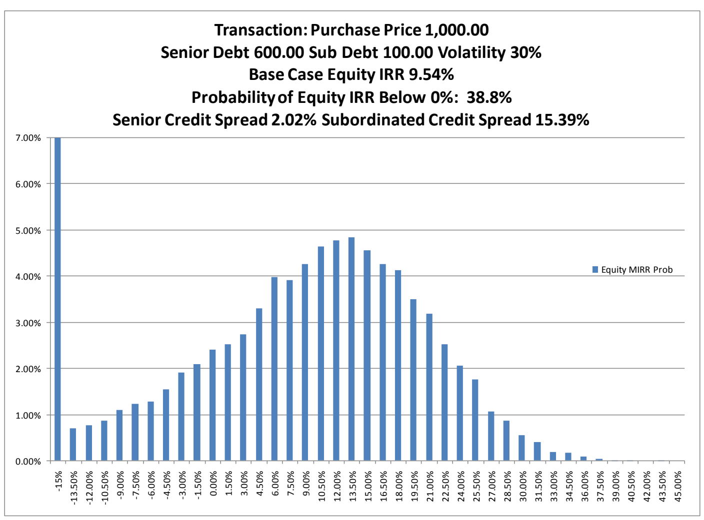
less or equal to zero is 3.2%. Without debt, the distribution is not skewed. The credit spreads are not shown on the graph because no debt financing is assumed.



To make the example a bit more interesting, assume the investment is financed with 600 of senior debt and 100 of subordinated debt. Without volatility, the base case IRR increases to 9.54%, but this statistic does not tell us about the risk to either debt holders or equity holders if this financing structure is implemented. Through running the Monte Carlo simulation, one can see that the probability of realizing a return below zero has increased to 24% and that there is a chance of earning as much as 27% rather than being limited to 16%. Further, the risk to the senior and subordinated debt holders can be measured. In this case, the senior debt needs a credit spread of .30% and the subordinated debt needs a credit spread of 3.37% to make the debt have equivalent returns to risk free debt. The large bar on the left of the graph measures the probability that the equity holders will loss there entire investment. Since the equity holders cannot lose more than their investment, the distribution of equity IRR is highly skewed, meaning that there is a limited downside and a wide variation of upside returns.



Let's say you are not very sure about the 20% volatility estimate. One can change the volatility and see what happens to the risks and returns of debt and equity holders. If the volatility of cash flows is 30% instead of 20% then the upside and the downside for equity holders increase (the scale of the graph below is different from the prior graph.) The probability of a return below zero increases to almost 40%, but the relative chances of earning a very high return also increase. The increased volatility has a dramatic effect on the required credit spreads to debt holders. The senior credit spread increases from .24% to 2.02% and the minimum required spread on subordinated debt jumps to 15.39%.



If you have never worked with Monte Carlo simulation, the statistics and pictures shown on the graphs may seem exciting. The analysis provides answers to things like minimum credit spreads and probability of losing money that would not be possible to compute with the judgmental analysis

described at the beginning of the chapter. Once you see how easy it is to implement you may become addicted to simulation and insist on including it in all of you future work. But before you get too enthusiastic, you should understand how easy it is to change a few variables and get dramatically different results.

To add this type of analysis to a financial model requires first reformulating equations from a financial model into time series equations and then evaluating the possible variability in selected output variables through using a random variable combined with a probability distribution and Monte Carlo simulation. The mechanical procedure for constructing the analysis is described in three sections below. The discussion begins with description of how to build a one column financial model used in the exercise. Next, the theory and mechanics of adding a simple time series equation into the financial model is addressed. The final section covers techniques which allow computation of risk statistics through making many different scenarios with different random numbers.

Step 1: Deterministic Financial Model

The first and perhaps the most difficult step in this exercise is not the time series equation or the Monte Carlo simulation, but rather creating a financial model that computes the default on different types of debt and the cash flow to equity. To make a model that measures default on multiple types of debt, a set of multiple equations which shows debt balances and cash flows is first presented. Once this model is established, a simple one column model that produces the same results is presented. This one column model is in turn used as the basis for the time series equations. In order to construct a simulation with both senior and subordinated debt, the figure below shows how create a model which first sweeps all cash to pay senior debt, then, once senior debt is paid, a similar cash flow sweep is used to pay subordinated debt and finally, once subordinated debt is repaid, the remaining cash flow goes to equity holders. To make a model with these cash sweeps, the debt balance for senior and subordinated debt are set-up as described in the last chapter and cash flow sub-totals are computed to test the ability to pay-off debt securities. The repayment of debt is computed by using all available cash flow less interest expense until all of the debt is repaid (the minimum of the opening balance or the cash flow less interest.) Once the cash flow after senior debt is computed, the same process is used for subordinated debt. There is a small complication in the subordinated debt calculation because interest must be capitalized until the cash flow after senior debt is positive. If the financial model is set-up in this manner, then the closing balance of the senior and the subordinated debt measures the default on debt. In the example below, there is no default on senior debt and a default of 208 on subordinated debt.

The present value of the default debt provides the basis for computing the minimum credit spread. In the Using a risk free rate for the senior debt and the subordinated debt (the addition of credit spreads is addressed in the next chapter), the financial model below illustrates how defaults and equity cash flow can be computed.

Inputs	
Operating	
Purchase Cost	1,000.00
Cash Flow	100.00
Volatility	20%
Financing	
Risk Free Rate	5%
Senior Debt	700.00
Subordinated Debt	200.00

Outputs					
MIRR to Equity	-100.00%		NPVo of Equity at Rf		($95.24)
MIRR to Project	2.32%		NPV of Project at Rf		($216.98)
		PV	Debt		FV
Default on Senior Debt		-	700.00		-
Default on Subordianted Debt		127.83	200.00		208.22

Model											
Year	0	1	2	3	4	5	6	7	8	9	10
Cash Flow		100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Senior Debt											
Opening Balance	-	700.00	635.00	566.75	495.09	419.84	340.83	257.88	170.77	79.31	-
Less: Repayment	**-**	**65.00**	**68.25**	**71.66**	**75.25**	**79.01**	**82.96**	**87.11**	**91.46**	**79.31**	**-**
Add Debt Issued	700.00										
Closing Balance	700.00	635.00	566.75	495.09	419.84	340.83	257.88	170.77	79.31	-	-
Interest Rate	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%
Interest Expense	**-**	**35.00**	**31.75**	**28.34**	**24.75**	**20.99**	**17.04**	**12.89**	**8.54**	**3.97**	**-**
Cash Flow After Senior Debt	-	-	-	-	-	-	-	-	-	16.73	100.00
Subordinated Debt											
Opening Balance	-	200.00	210.00	220.50	231.53	243.10	255.26	268.02	281.42	295.49	293.54
Add: Interest Capitalised	-	10.00	10.50	11.03	11.58	12.16	12.76	13.40	14.07	-	-
Less: Repayment	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**1.95**	**85.32**
Add: Debt Issued	200.00										
Closing Balance	200.00	210.00	220.50	231.53	243.10	255.26	268.02	281.42	295.49	293.54	208.22
Interest Rate	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%
Interest Expense	0.00	10.00	10.50	11.03	11.58	12.16	12.76	13.40	14.07	14.77	14.68
Interest Capitalised	0.00	10.00	10.50	11.03	11.58	12.16	12.76	13.40	14.07	0.00	0.00
Interest Paid	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**14.77**	**14.68**
Cash Flow to Equity	-100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Free Cash Flow	-1,000.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

This model that contains about twenty lines can be reduced to a single line (column). This process involves the following four step process:

- First compute the future value of cash flows with no financing, showing the year by year cash flow in a single column over the 10 years;
- Second, compute the future value of the operating cash flows as well as the future value of the senior debt and the subordinated debt and place the values below the future value of the operating cash flows;
- Third, compare the future value of the operating cash flows to the future value of senior and subordinated debt. If the value of the future cash flows is less than the future value of the senior debt, then there is a default on the senior debt (equivalent to the debt outstanding at the end of the period in the above diagram.) The default on the subordinated debt is the future value of cash flows less the senior and subordinated debt, plus the default on the senior debt.
- Fourth, compute the cash flow to equity as the future value of cash flow with no financing less the future value of senior and subordinated debt. The modified internal rate of return on equity can be computed as the future value of equity cash flow divided by the equity investment raised to one divided by the number of years.

An example of this one column model that produces the same results as the above multi-line model is shown below.

Year	FV Factor	Operating Cash Flow
0	1.63	(1,000)
1	1.55	100.00
2	1.48	100.00
3	1.41	100.00
4	1.34	100.00
5	1.28	100.00
6	1.22	100.00
7	1.16	100.00
8	1.10	100.00
9	1.05	100.00
10	1.00	100.00

FV of Operating Flows	1,257.79
FV Factor	1.63
FV of Senior Debt	1,140.23
FV of Sub Debt	325.78
Default on Senior	-
Default on Sub	208.22
Equity	-
Project MIRR	2.3%
Equity NPV	(95.24)
Project NPV	(216.98)
Equity MIRR	-100.00%

Step 2: Time Series Equations

To incorporate mathematical analysis such as probability of achieving a return below the risk free rate, the operating cash flow must be modeled as a random variable driven by the volatility parameter. The most basic way to create such a random is to assume the cash flow follows the well known random walk process. In a random walk, the variable (in this case cash flow) can move up or down from one period to the next with equal probability. One can think of this process by imagining a drunken man starting to walk along a line. After each step, the man may stumble in one direction or another with equal probability. After he takes the first stumble, the process begins again from the point of the last stumble, and the man can stumble again in each direction with equal probability. The prior stumble has no effect on the direction of subsequent stumbles. Depending on the how long many steps the man takes, the drunken man may wander in quite large directions from the initial starting point. The range in size of each stumble can be though of as volatility. The random walk model can be described with the following equation of the current price purely a function of last periods price and random movement -- ε:

$$P_t = P_{t-1} + \varepsilon.$$

In the above equation, ε is a random term that can move up or down with equal probability and movements of ε in one period are independent of movements in other periods. This term can be replaced by the volatility percentage combined with a random number draw as shown in the formula below:

$$P_t = P_{t-1} + \text{Random Number x Volatility Percentage x } P_{t-1}.$$

In the above equation, if the volatility parameter is 20% and the random number happens to be 1.0, then the next price is 20% above the earlier price (the drunk man would move 20% to the right.) If the

volatility parameter were zero, there would be no movement in prices no mater what the random number. Finally, if the random number happens to be zero, then the price does not move no matter what is the volatility parameter. In this simple equation, the random number must have an expected value of zero (i.e. when a lot of random numbers are drawn, the average of the random numbers should average zero.) The current period price is a function of the prior period price and the random variable. Volatility influences the equation because it magnifies the effect of the random shock.

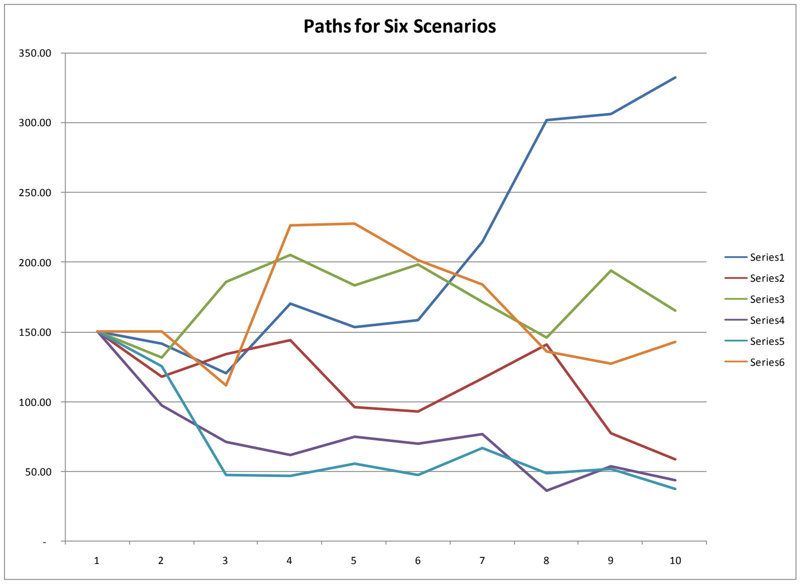
To mechanically implement the time series equation, assume that the volatility is the percent change derived from a normal distribution. In this case the random number should range from about -4 to +4 reflecting the potential number of standard deviations away from the average -- the most extreme cases are where the price moves by four times the standard deviation of 20%. To create the time series in excel with a normal distribution, the functions RAND() and NORMSINV can be used. The NORMSINV function accepts a number from zero to one and produces the standard normal value that reflects the possible range in the percent change. Since the RAND function generates a number between zero and one, this function can be used to derive a random draw from a normal distribution as shown in the formula below.

$$\text{Cash Flow}_t = \text{Cash Flow}_{t-1} + \text{Volatility Parameter x (NORMSINV(RAND()) x Cash Flow}_{t-1}$$

To implement this formula, simply change the formula in the second year after the first number is entered (i.e. the second 100 in the above diagram.) Instead of making the second cash flow equal to the first cash flow, add a term that allows the cash flow to vary depending on a random draw. The revision involves making the second cash flow equal to the first cash flow plus the first cash flow multiplied by the volatility and NORMSINV(RAND()). Once this formula is established, for the second year cash flow, copy the same formula to all of the remaining rows. This means that a new random number will be drawn for each year after the first year.

Step 3: Monte Carlo Analysis

Once the series of cash flows is established for one column (for the ten years there are nine different random draws), Monte Carlo simulation simply means the same column is copied over and over again creating many different scenarios with different random numbers. In excel 2007, there are more than 16,000 columns meaning that using the SHIFT, CNTL, → and CNTL, R technique (documented in Chapter 2) to copy formulas across all of the columns results in more than 16,000 different scenarios, each scenario having nine different random numbers. The first six cash flow scenarios assuming 30% volatility are shown on the chart below. Note how some scenarios result in a cash flow near zero while others result in very high cash flows. Further observe that in general, the variance in prices increases as time passes.



The instructions in the above paragraph involving copying the time series equation gets the Monte Carlo job done. The things that were done to create a simulation are described in a more formal way below. Applying Monte Carlo simulation to the equation above involves a few steps where a time series equation is defined, random numbers are generated, a path of prices is computed and the process is repeated for multiple scenarios. The following six step procedure describes the general process for creating a Monte Carlo simulation:

Step 1: Choose the time increment (the period could be a day, a month or a year), the number of periods for each path, and the number of simulated price paths.

Step 2: Begin the process with the current price and draw a random number between zero and one to project how prices in the first period will change from the first period to the second period.

Step 3: Convert the random number into a standard normal draw through filtering the random number through a normal or another distribution. This means using the inverse of the normal distribution as the normal probability distribution defines the probability of a number such as zero (the mean) and plus or minus one (the standard deviation). If the filter is the normal distribution, random number draws between 0 and 1 result in filtered numbers between -4 and 4 with a mean of zero and a standard deviation of 1.

Step 4: Use the filtered random number – the standard normal value -- to compute the next period price in the time series equation.

Step 5: Repeat the process of drawing a random number, filtering it through a probability distribution and computing a new price for subsequent periods in the first scenario.

Step 6: Once the process is completed for the total number of periods for one price path, begin the same process for the next simulation and continue to repeat the procedure for the total number of simulations.

To illustrate how this Monte Carlo simulation process works, assume a cash flow forecast is made for 10 years into the future with a volatility of 20%, a starting cash flow of 100 and no mean reversion or time trend. The simulation begins by making a random draw for the first month and the first cash flow scenario, say the draw results in a number of 0.6. Using the inverse of a standard normal distribution, the value of 0.6 translates into a standard normal value of .25 (a random number of 0.5 would result in a standard normal value of zero.) This standard normal value is then multiplied by the volatility of 20% yielding a percent change in cash flow of .25 x 20% or 5% from the first year to the second year. The cash flow in the second year therefore becomes 100 x (1+0.05) or 105. To model cash flow in the third year for this first scenario, the same process is used – drawing a random number, filtering it through an inverse normal distribution and adjusting for volatility -- but the cash flow that is the starting point for the formula is 105 instead of 100. After cash flow in the third year is established, another random draw will be made resulting in a new cash flow for the fourth period. By the 10th year, a price series is created that reflects the impacts of making 9 random draws. Once the entire first cash flow scenario is established, the same process is used to generate a second set of cash flow for 10 years. Ultimately, the procedure for computing cash flow paths is repeated many times for the 10 year period – often making 1,000 to 100,000 price simulations. Once all the cash flow paths are simulated, the distribution of cash flow from the 20^{th} month can be analyzed through computing the standard deviation and other statistics.

Risk statistics such as the required credit spread and the probability of achieving a return below zero can be easily computed once the Monte Carlo simulation is complete. The minimum credit spread is simply average of the amount of default across all of the scenarios divided by the debt outstanding. The probability of achieving a return below zero involves first computing a probability distribution of returns as described above using the FREQUENCY function and then using the MATCH and INDEX function to find the probability corresponding to a return of zero.

This model would be inappropriate for use in real world analysis because it assumes a normal distribution of random numbers; no mean reversion or price boundaries and no interdependence between different variables. Despite the simplifications, the exercise does demonstrate how a volatility parameter drives various aspects of risk analysis. Actual Monte Carlo simulations are more complex than this simple example because of the time series equations are include added parameters (trends, boundaries, jumps and so forth), because of transformation to logs, and because the random draws may be filtered through some other type of probability distribution. In the more complex case, the time series equation (without jumps, price boundaries or correlations and not in logs) can be written as:

$$\text{Price}_t = \text{Price}_{t-1} \text{ x Trend Factor} + (\text{Price}_{t-1} - \text{Average Price adjusted for Trend}) \text{ x Mean Reversion Factor} + \text{Volatility Parameter x Inverse Normal Distribution (Random Draw) x Price}_{t-1}$$

While this equation is more complex than the simple equation, no matter how the time series equation is specified and no matter what probability distribution is used, the process always boils down to the idea of drawing a random number, adjusting it for volatility and making the next period price a function of the prior period price.

Alternative Types of Time Series Equations

The next paragraphs discuss more complex aspects of time series models that can be used in forecasting prices, demand and other factors. The discussion moves from simple time series models that include only a volatility parameter to more complex models with time trends, mean reversion, lower

and upper boundaries, jump processes and correlations. In this section, parameters such as volatility, mean reversion factors, correlations and other factors are assumed to be given.

Brownian Motion and Normal Distributions

The most pervasive set of time series models applied to financial instruments such as stocks, options and other derivatives are random walk models built on assumption that rates of change follow a normal distribution and that changes in price are independent of past prices. For example, in the famous Black-Scholes model discussed in Chapter five an assumption is that prices follow a normal distribution and historic price changes have nothing to do with prospective price changes. Given that Brownian Motion underlies many applications in finance, a few characteristics of this time series model are addressed below, including whether real world data confirms to implicit assumptions. The most fundamental property of these models is that prices are non-stationary, meaning that they can wander all about without coming back to some average level.

The theory of efficient markets is consistent with the notion that stock prices follow a random walk process (price changes should be independent with one another although stock price changes do not have to be normally distributed.) The basis of efficient markets theory is that any change in stock prices arises from new information – not earlier information which has already been incorporated in the price. This means that future price changes have nothing to do with price changes that occurred in earlier periods. If new information is normally distributed, and prior information has already been included in prices, the price process follows a random walk process and where past prices -- $P_{t-2,}$ $P_{t-3,}$ and so forth have no forecasting value. If past prices are irrelevant, there is no mean reversion in prices. While this theory may apply to stock prices, it clearly does not apply to many variables that are entered into financial models and that drive risk assessment. For example, in the case of electricity price, a change of one direction in the price of electricity often means that subsequent prices will change in the opposite direction. In this situation past electricity prices are clearly very relevant in predicting future price changes suggesting that electricity prices do not follow a random walk.

The common equation used in financial economics is similar to the equation used in the simple example in the last section in which random draws are extracted from a normal distribution. However, there is slight technical difference involving transforming the data to logarithms. This difference (which many technicians like to emphasize to show how smart they are) involves the assumption of continual changes rather than discrete changes. Here, the price in the next period is not defined as this period's price multiplied by one plus the volatility as in the above equation, but rather this price is the exponent of volatility as shown below.

Discrete:	$P_t = P_{t-1}$ * (1+ P_{t-1} x Volatility Percent x Standard Normal Draw)
Continuous:	$P_t = P_{t-1}$ * (exp(Volatility Percent x Standard Normal Draw))

If the time increments are continuous rather than discrete, the random walk becomes Brownian motion. In Brownian motion, the ε term is a draw from a normal distribution with a variance that increases on a linear basis over time. Use of the continuous distribution prevents the possibility of a negative price (the exponent cannot be negative) and involves somewhat more complicated mathematical equations. Rather than burdening you with these more complex formulas which do not really affect the most of the fundamental issues, the continual distribution is discussed in Appendix A at the end of the chapter.

One of the basic questions before deciding to implement a model with Brownian motion is determining whether the underlying assumption of a normal distribution and independence of price changes is applicable. The question of whether rates of change are normal can be addressed by inspecting a frequency graph of percent changes, while a simple and crude method to test whether there is no mean reversion in the series is to test whether rates of change over longer increments have higher volatility than shorter periods amplified by a factor of the square root of the number of periods.

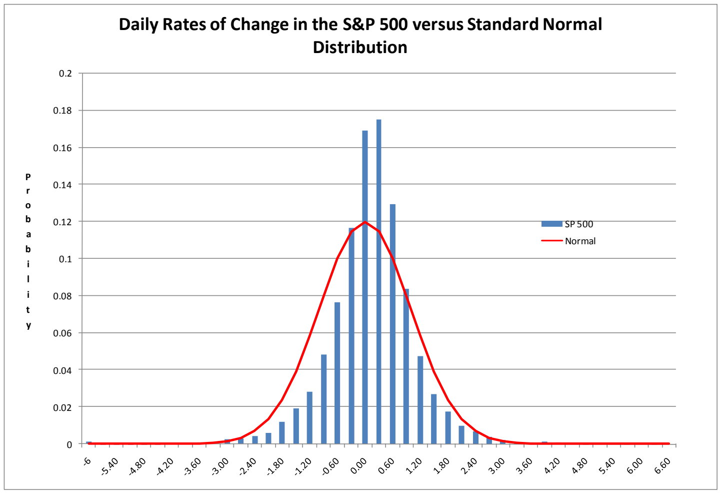
In testing whether rates of change follow a normal distribution, the two functions discussed above – NORMDIST which returns the probability and the NORMSDIST, which returns the value of a distribution there are two functions in excel that help one to understand whether a distribution is normal and thus whether it is appropriate to apply draws from an standard normal distribution in a Monte Carlo analysis.

In computing the distribution of a series of percent changes compared to the percent changes that would come from a normal distribution, the following steps can be followed:

- Compute the percent change in the series and the mean and the standard deviation of the series of percent changes
- Compute a standardized distribution by subtracting the mean percent change from each of the observations and divide by the standard deviation.
- Create bins for a frequency distribution (begin with something like -6 or -4).
- Create a frequency distribution and the probability of the series of percent changes (divide the frequency by the total number of observations to convert the frequency distribution into a probability distribution.
- Compute the probability distribution for a normal distribution using the NORMSDIST function with a switch for a non-cumulative distribution.
- Use the handy F11 key to make a graph of the percent changes relative to the normal distribution.

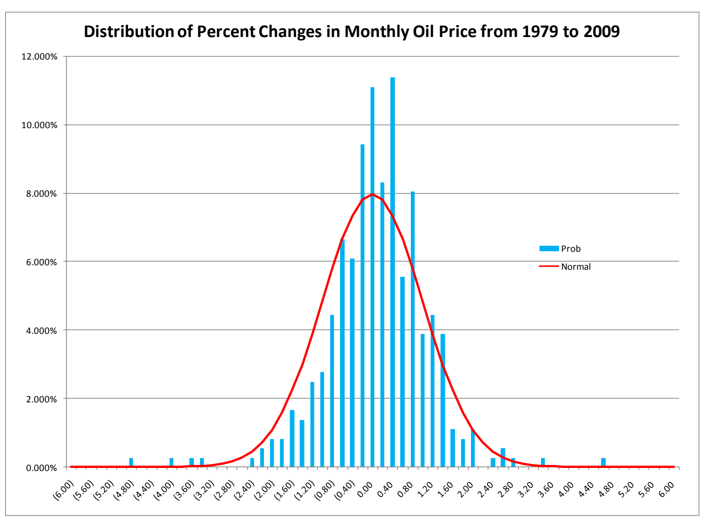
Once you get used to this process, it will go very quickly. The results for daily changes in the value of the S&P 500 is shown in the graph below.[6] Many studies have been made regarding whether stock price changes follow a normal distribution. Without making any highly sophisticated econometric analysis, the graph below illustrates that daily stock price changes have not been normal.

6 All of the data and methods are included on the CD.



Since changes in stock price do not follow a normal distribution, one could attempt to fit another distribution, but that would be difficult and dangerous. The danger comes from the very small chance that the percent change is price could be six times away from the mean. Since 1950, there have been 17 days that have been in the negative six bucket. In a normal distribution, the chance of reaching minus six standard deviations away from the mean is one day in 4.02 million years (this can be computed using the NORMSDIST function, plugging in -6 and then multiplying the probability by 252 days per year).

The story is similar for the distribution of oil prices. The graph below shows the distribution of the monthly rate of change in real oil prices compared to a normal distribution. As with the S&P 500, there are a number of observations that exceed four standard deviations and would be almost impossible if monthly oil price changes were really normally distributed.



The distribution of prices can be used to evaluate whether mean reversion or other constraints on movement of prices is present. If the data came from Brownian motion, then the standard deviation of prices computed on the basis of average annual data would be greater than the daily volatility by a factor of the square root of 252 (the number of trading days in a year.) Similarly, the standard deviation of percent changes computed on average monthly prices would be the square root of 12 or 3.46 multiplied by the standard deviation. The table below shows the relationship between the annual standard deviation and the daily standard deviation of percent changes in price for oil prices and the S&P 500 index. In the case of oil prices, the lower volatility for longer periods implies that mean reversion exits while in the case of the S&P 500, there is less evidence as the daily volatility is not much more than the monthly or the annual volatility.

S&P 500 Volatility			
	Daily	Monthly	Annual
Std Dev	0.96%	3.55%	13.80%
Adjusted	15.22%	12.30%	13.80%

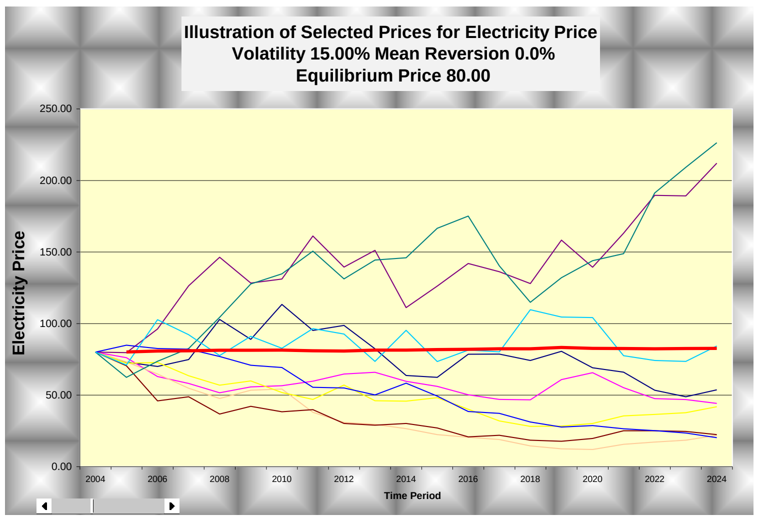
Percent Change in Oil Price			
	Daily	Monthly	Annual
Standard Deviation	2.55%	8.21%	23.90%
Adjusted	40.44%	28.43%	23.90%

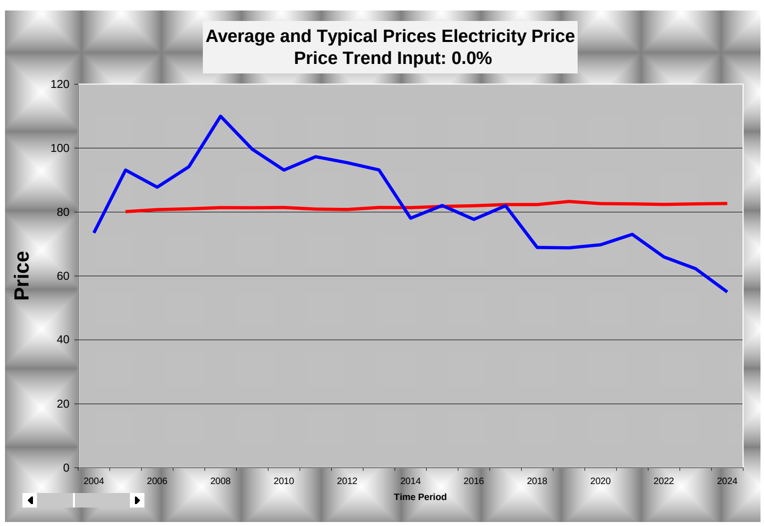
Volatility IRR’s and Probability of Default

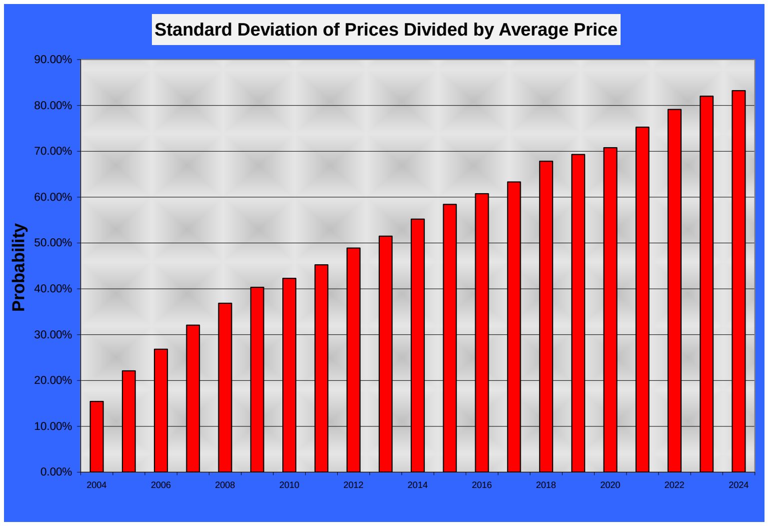
In describing how various issues such as how mean reversion, price trends, correlations, price boundaries and price jumps affect time series analysis and ultimately risk measurement, a relatively simple model of an electricity plant is used to illustrate the effect of different parameters. The plant is assumed to earn revenues from an uncertain price and capacity factor and it has to purchase natural gas for generation. It is assumed to have a cost of 500 million with 200 senior debt, 50 of subordinated debt and 250 of equity. The initial exercise assumes the price of electricity comes from a Brownian Motion process with different volatility parameters of 15%, 25% and 40%. Variables such as the capacity factor and the natural gas price are assumed to have no volatility. This case is similar to the simple case discussed in the earlier section as there is no mean reversion, price boundaries or other time series parameters other than volatility except that a log-normal distribution is assumed (as described in Appendix A.)

The parameters required to run a simulation for a random walk include the starting price, the volatility, the number of periods, the number of simulations and whether the data is transformed to logs. To illustrate this process, the exercise assumes beginning price of 80, volatility of 15%, 25% and 40% and 20 annual periods for the analysis. 1,000 simulations have been run to achieve a distribution of possible prices. The first graph below shows results of a few of the price simulations in the case with volatility of 15% while the second graph shows the median simulation and demonstrates a typical movement in price over time. The third graph divides the standard deviation of prices in the various simulations by the average price across the simulations. Even though the volatility is only 15%, the first graph shows that the price can eventually move from 80 to 200 while in the low cases the price is near zero. As expected, the variation of prices increases through time. When all of the simulations are

averaged for each period, the price is just about equal to the beginning price as expected (software used to compute Monte Carlo simulation along with a set of instructions is provided on the CD that accompanies this text.) Finally, as expected with Brownian motion, the standard deviation increases as the length of the period increases.







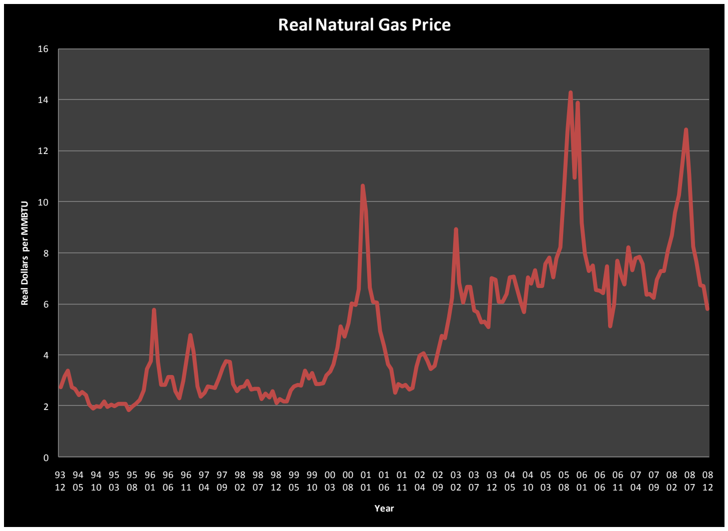
From a financial risk perspective, results for the Brownian motion case establish a benchmark for which some of the subsequent analysis can be compared. A variety of financial statistics that gauge the risk of equity and debt can be computed from the simple financial model as long as the model includes provision for default on debt. The table below shows that if the volatility of electricity prices is 25%, that the minimum required credit spread for senior debt is 3.2% and the probability of a equity IRR below zero is 34%. In the lower volatility case the minimum required credit spread is only .35% while the higher volatility case results in a high required minimum credit spread of 3.2%.

	Volatilty	Mean Reversion	Price Boundaries	Jump Process	Correlations	Probability of Total Equity Loss	Probability of Loss on Senior Debt	Probability of Loss on Subordinated Debt	Median Equity IRR	Median Project IRR	Probability of Negative Equity IRR	Probability of Negative Project IRR
Case 1	25%	None	None	None	None	21%	3.20%	13.30%	9.90%	7.70%	34%	26%
Case 2	15%	None	None	None	None	8%	0.35%	2.95%	12.20%	9.10%	22%	13%
Case 3	40%	None	None	None	None	32%	8.47%	24.89%	5.17%	5.05%	45%	37%

Mean Reversion and Long-Run Equilibrium

Variables which move according to Brownian motion wander around and gradually move more and more away from the initial value. Many, if not most economic variables move in cycles rather than in random walks. In the case of price variables this is due to the simple fact that high prices will prompt increased supply thereby moderating the price increase while very low prices will cause new supply to slow down. For example, most commodity prices including oil, gas, coal and electricity should eventually move in the direction of their long-run cost of production rather than continually moving up or down without limits. This tendency of oil prices is noted by Dixit and Pindyck as follows: "…while in the short-run the price of oil might fluctuate randomly up and down, in the long-run it ought to be drawn back towards the marginal cost of producing oil."[7] The propensity to move to equilibrium amounts is known as mean reversion. The tendency for natural gas prices to gradually revert to average levels is illustrated in the graph below where prices seem to increase for a couple of years and then fall back.

[7] Ibid, Dixtit and Pindyck, page 74.



From a statistical standpoint, the mean reversion in the above graph can be demonstrated by computing the standard deviation of percent change in price for different time periods. For example, the average price can be computed for every three years. Then, the standard deviation of the percent change in three year average prices can be compared to the standard deviation of price changes computed on a monthly basis. Without mean reversion, the standard deviation of percent changes for the longer three year period should be six times as great as the monthly standard deviation (the square root of 3 x 12.) In fact, the standard deviation of percent changes for the longer period three year period is smaller than the standard deviation of the monthly percent changes as shown in the table below.

Natural Gas Percent Changes			
	Montly	**Annual**	**3-Year**
Standard Deviation	14.93%	28.29%	9.25%
Adjusted	51.71%	28.29%	5.34%

For electricity prices, mean reversion is a particularly prominent statistical feature over both short-term and long-term periods. The mean reversion of electricity contrasts starkly with stock prices, which supposedly have the property that historic prices do not influence future price changes. Stocks store the value of a company created by cumulative past decisions and past events and in theory, they change only when new information arises that was is not already reflected in the current price. Electricity, on the other hand, cannot be stored meaning that once an event occurs that affects prices such as hot weather or a plant outage, the next period price starts over at production costs with normal weather or plants back in service. Electricity prices are the polar opposite of stock prices in terms of the relevance of past prices and the tendency to revert to mean levels.

The reason mean reversion is often present in commodity prices is because of the way new capacity is developed when prices move. In many industries including real estate, telecommunication, infrastructure and commodities, the behavior of developers is somewhat unpredictable and it arguably may even irrational in the short-term. However, at some point construction slows or accelerates

because of the prices. This behavior forces prices to move in the direction of long-run marginal cost. The real world activities of developers compared to the theoretically rational behavior assumed in many pricing models is made by Standard and Poor's as follows.[8]

> Generally most models assume rational behavior by market participants, and that a system will operate economically in the most efficient manner, given the physical constraints of the system. Unfortunately, market participants do not always behave rationally or uniformly – fortunately they do not behave recklessly either…Underlying every model is the assumption that the markets will operate in a perfect long-term equilibrium, and that they will price electricity at the marginal cost of production. Unfortunately the reality of energy markets appears much different. Equilibrium lasts momentarily at best. A sudden change in fuel prices, a change in system load, or the presence of a new power station will upset the equilibrium.

The behavior described in the above quote can be modeled as a random process that gradually moves back to equilibrium levels driven by production costs, but does not remain at stable levels. The fact that new developers of a project cannot indefinitely lose money (because no new investments would occur and reduced supply would push prices up) or continually earn supernormal economic profits (because of new entrants would be attracted to the market pushing prices down) is represented by movement of prices to the long-run marginal cost. To create a mathematical model of economic behavior that pushes prices move to long-run marginal cost, a mean reversion factor should be included in time series analysis.

When an economic variable reverts to average levels, the variable can be modeled as a mean reverting process rather than as a random walk process. For time series equations with mean reversion, variables still have random shocks driven by volatility; but after a random shock causes the process to move away from mean levels, prices in subsequent periods tend to move back to a defined average. The speed at which prices revert to the long-run equilibrium after a random shock is generally expressed in terms of annual percentage amounts. For example, if the long-term equilibrium is 100 and a shock causes prices to move to 150, then a mean reversion factor of 40% would mean that the next year price tends to move down by 50 x 40% or 20. If the prices are expressed in annual terms, one divided by the mean reversion factor can be thought of as the amount of time it takes to come back to the equilibrium level. The basic equation for a mean reverting process is:

$$P_t = P_{t-1} + \text{Mean Reversion Factor} \times (P_m - P_{t-1}) + \varepsilon$$

In the above formula, the mean price term P_m may approximate the long-run cost of production and the ε term incorporates volatility and draws from a normal distribution as was the case for random walk processes. The mean reversion factor in the equation is generally between zero and one. If the mean reversion factor is 1.0, then the P_{t-1} terms cancel and the equation becomes $P_t = P_m + \varepsilon$ implying that prior period observations have no value in forecasting; everything starts again from the average. In this case, there are movements away from the mean level due to random shocks, but the process starts over at the mean level in the next period – one could think of the weather as following a process like this. On the other hand, if the mean reversion factor is zero, the equation is the same as the random walk processes described above where prior period prices have no forecasting value and the process is non-stationary. One can think of the mean reversion factor in terms of how many years it takes for a

[8] Rigby, Peter N., "Risks from Left Field: Is There a Problem with Pricing Models," Comments at the 15th Annual Global Power Markets Conference, April 16-18, 2000.

market to reach equilibrium after a shock. For example, if a market is constrained with high prices and it takes three years to construct new plants, then the mean reversion factor would be 33%. Alternatively if there is excess supply and demand growth would cause take ten years until a supply/demand balance occurred, then the mean reversion factor would be one divided by ten or 10%.

The equilibrium price or long run marginal cost can be modeled as a stochastic variable resulting in a two factor model. To implement such a two factor model, two random draws are made and a volatility parameter applicable to the equilibrium price is included in the time series model. The long-term marginal cost or equilibrium price that is input in a time series model can also vary from period to period. If a forward market exists for a commodity (the case for many electricity markets, natural gas markets and oil markets), the mean price can be modeled using the forward price. In general time series models should use all available market data. Random shocks can cause spot prices to deviate from the forward prices, but prices move back to the long-run level after the effect of the shock has worn off.

Mean Reversion IRR's and Probability of Default

When applying time series analysis and Month Carlo simulation to investments, mean reversion can have a dramatic effect on the various risk measures. To illustrate this, consider the natural gas generating plant in the example introduced above. In addition to the volatility parameter, various different mean reversion factors are included in the simulation. As in simple example introduced above and in the Brownian motion case, the first step is constructing the time series equation, the second step is running a Monte Carlo simulation to obtain scenarios with different prices, and the final step is using the different prices in the financial model to obtain a distribution of IRR's, minimum credit spreads or other financial variables.

Simulation of prices that include a mean reversion parameter demonstrate differences in the characteristics of a mean reverting time series from a random walk series. To illustrate price patterns that result from a mean reverting time series, case have been developed which volatility of 25% and 40% as in the last case, but also mean reversion factors of 10%, 30% and 50%. The exercise of simulating prices with mean reversion allows one to inspect the simulation results and evaluate whether the price patterns reasonably represent possible movements in the price of electricity. Outcomes of the case with mean reversion of 30% are presented on the three graphs below. The first graph illustrates that with mean reversion, the extreme price cases are far more unlikely. Unlike the case of a random walk process, the standard deviation of prices in a series with mean reversion does not increase over time.

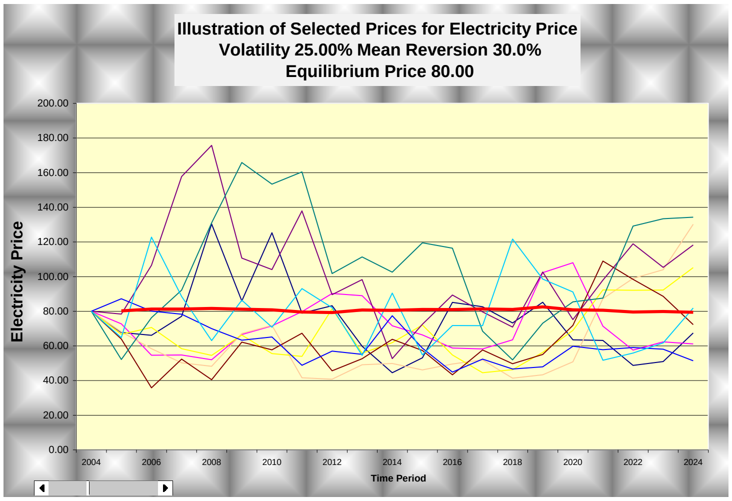
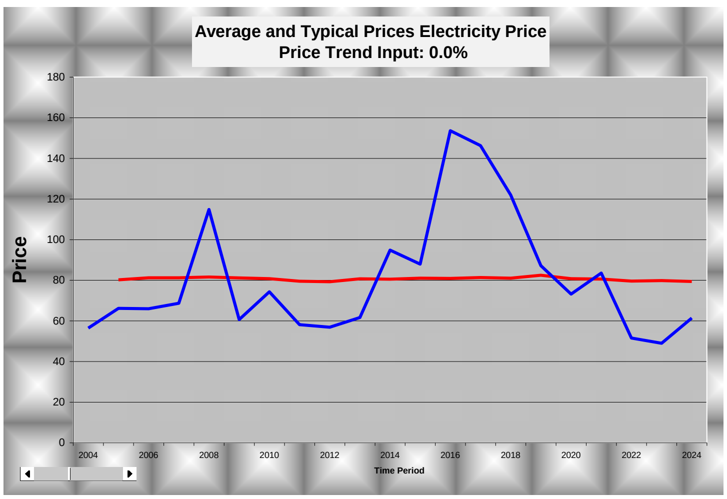
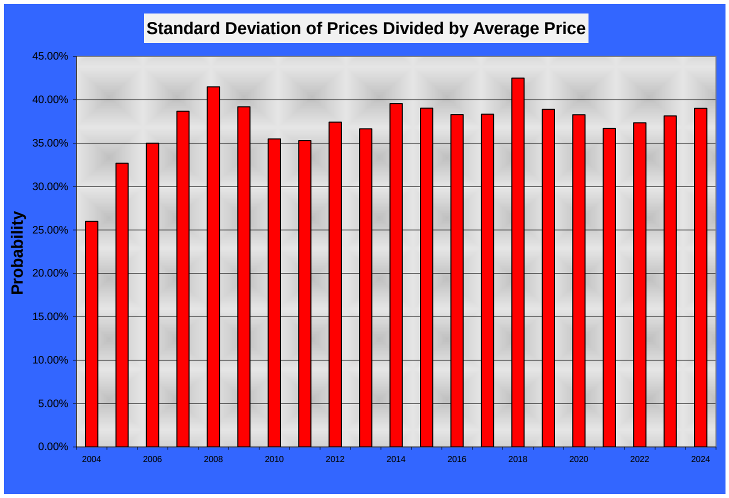


Illustration of Selected Prices for Electricity Price
Volatility 25.00% Mean Reversion 30.0%
Equilibrium Price 80.00
Electricity Price
200.00
180.00
160.00
140.00
120.00
100.00
80.00
60.00
40.00
20.00
0.00
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
2024
Time Period


Average and Typical Prices Electricity Price
Price Trend Input: 0.0%
Price
180
160
140
120
100
80
60
40
20
0
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
2024
Time Period



Adding mean reversion factors dramatically changes the measures of risk relative to the cases without mean reversion. Even if a parameter of only 10% is assumed -- which assumes that it takes 10 years to reach equilibrium through construction of demand increases -- risk is reduced dramatically as shown by the senior debt probability of loss on the last line of the table. Without mean reversion, the minimum required credit spread is 8.47%. In the case with 10% mean reversion, the minimum required credit spread decreases to only 1.82%. If mean reversion is 50% implying three years to reach equilibrium, the risk for senior lenders or subordinated lenders is to almost zero. Clearly the presence of mean reversion has a dramatic effect on the risk analysis. It should be clear from the table below that ignoring mean reversion or over-estimating mean reversion can lead to big mistakes in evaluation of the risk of an investment.

	Volatilty	Mean Reversion	Price Boundaries	Jump Process	Correlations	Probability of Total Equity Loss	Probability of Loss on Senior Debt	Probability of Loss on Subordinated Debt	Median Equity IRR	Median Project IRR	Probability of Negative Equity IRR	Probability of Negative Project IRR
Case 1	25%	None	None	None	None	21%	3.20%	13.30%	9.90%	7.70%	34%	26%
Case 2	15%	None	None	None	None	8%	0.35%	2.95%	12.20%	9.10%	22%	13%
Case 3	40%	None	None	None	None	32%	8.47%	24.89%	5.17%	5.05%	45%	37%
Case 4	25%	50%	None	None	None	0%	0.00%	0.00%	12.78%	9.50%	2%	0%
Case 5	25%	30%	None	None	None	2%	0.00%	0.00%	12.20%	9.13%	12%	3%
Case 6	25%	10%	None	None	None	14%	0.23%	2.13%	10.74%	8.21%	28%	18%
Case 7	40%	50%	None	None	None	2%	0.00%	0.00%	11.73%	8.83%	12%	3%
Case 8	40%	30%	None	None	None	10%	0.00%	0.40%	10.82%	7.83%	24%	14%
Case 9	40%	10%	None	None	None	25%	1.82%	7.60%	7.55%	6.34%	39%	30%

Time Series Equations with Trend Terms and Movement to Long-term Marginal Cost

In financial analysis of long-term investments, the long-term trend in a variable is likely to be a more important consideration than the short-term fluctuations around the trend. Studying the long-term price trends through evaluation of long-term marginal costs and potential changes in productivity and evaluating demand analyses is essential in valuation analysis. The issue of developing price forecasts for valuation is addressed in other chapters. This section only describes how to incorporate long-term trends into time series analysis.

The simple random walk equation above can easily be extended to include trends in expected prices so that the expected price change is not zero. Trends can be incorporated in the equation through adding a drift term to current prices or through incorporating explicit forward price forecasts. The simplest way to add information about future expected directions in prices is to include a trend term to the random walk process. Differences in the equation that arise from an assumption of continual compounding is assumed is shown in appendix A where the trend factor is part of the exponent.

$$P_t = P_{t-1} + \alpha * P_{t-1} + \varepsilon.$$

The "α" term in the above equation is the percentage growth rate in prices and the ε term incorporates the volatility parameter and the distribution assumption as with the previous equation. The economic rational for adding a trend term in the equation could be inflation (if prices are to be forecast in nominal terms), expected capital gains from retaining profits (if the price is a stock price), or changes in industry productivity (if long-term real prices are projected.) If explicit forward forecasts are used, the α term varies by year.

Productivity is defined as the level of inputs needed to produce a level of output. For many things, one would expect technological improvements to result in gradual increases in productivity and reductions in the real prices over a long period of time because humans should find ways of using inputs somewhat more efficiently. Unless there are constraints on existing resources (such as oil and gas) or new regulatory requirements, the idea is that one can always employ the same technologies and management systems as those that exist today. If there are changes in management systems and/or technological improvements in the use of resources, these things improve productivity. With increased productivity in a competitive industry, prices should decline over time.

In modeling trend terms related to productivity, inflation, capital gains or other factors, it is reasonable to expect that the trend term itself is a stochastic variable (implying the growth rate variable is not known with certainty). If the trend term is stochastic, a volatility parameter related to the distribution of potential trend terms can be incorporated in the time series equation. Here, where a volatility parameter is included in the trend term, the trend term is modeled with the same process as the random walk equation above. Cases in which volatility is included both in the trend term and the current value can be termed a two factor process. To include a stochastic trend term, the trend term in the above equation is replaced with a trend term defined by a second random draw as shown in the formula below.

$$\alpha_t = \alpha_{t-1} + \text{Standard Normal}_1 \times \text{Trend Volatility} \times \alpha_{t-1}$$

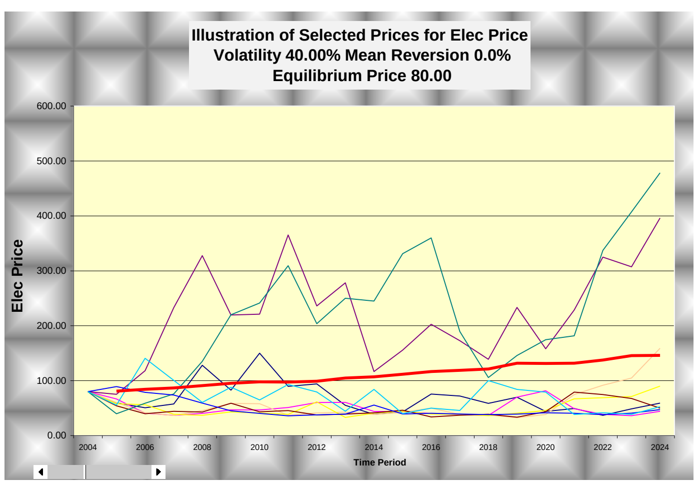
In the case of electricity prices, the volatility in short-term prices is driven by variability in weather, variation in hydro production and uncertainty in maintenance outages. On the other hand, volatility in the trend factor is driven by variation in future productivity that in turn is influenced by changes in the real cost of new capacity the heat rate of new plants. The volatility in short-term prices can be measured or simulated on an objective basis, but the volatility in trend factors is more difficult to estimate.

Price Boundaries and Short-run Marginal Cost

Even after carefully considering volatility, time trends, and mean reversion, prices and other economic variables modeled by a time series process can move to unrealistic levels because of the manner in which random numbers are drawn. For example, as noted above, if the prices can become negative in

models if discrete rather than continual changes are assumed. In time series equations with low mean reversion and high volatility, a time series equation can allow prices to become very high or very low for extended periods of time. One problem with these models is that extreme values prompt responses by consumers or suppliers that limit the extreme levels. Low prices tend to be limited by producers ceasing production and high prices tend to be limited by consumers who reduce demand. If prices fall below short run marginal costs, companies will chose not to produce. The reduced supply increases price and limits the decline in price to short-run marginal cost. This implies that prices in time series models should have a lower boundary defined by short-run marginal cost. Upper price boundaries can also be appropriate in time series models because if prices reach high levels, demand may be curtailed by factors such as the operating cost of alternative capacity.

Boundary conditions can be easily incorporated in time series models through limiting the downward or upward movement. The models can simply include a conditional factor that when prices drift above the upper boundary or below the lower boundary, the price is set to the boundary level. As with the other components of a time series model, the boundaries themselves can be stochastic variables with volatility and mean reversion. The effect of price boundaries depends on the level of volatility and mean reversion parameters which determines the likelihood of reaching the boundaries. To illustrate the effect of price boundaries, the natural gas plant case introduced above is adjusted to incorporate lower and upper boundaries. The case with 40% volatility is adjusted to include lower boundaries because with mean reversion there is a very small probability of default. A lower boundary of 42 is imposed which represents the marginal cost of running the plant is assumed along with a volatility of 10% applied to the boundary. The graphs that show the effect of price boundaries are presented below in the case with no mean reversion. The lower boundary condition with high volatility results in many of the price paths remaining near the lower boundary while some increase to very high levels. With high volatility and tight boundaries, the prices will concentrate around the boundaries.



The effect of lower boundaries on measures of financial risk is illustrated on the table below. In the case with no mean reversion, the minimum required credit spread for senior debt decreases from 8.47% to .72% just by adding the lower boundary. The

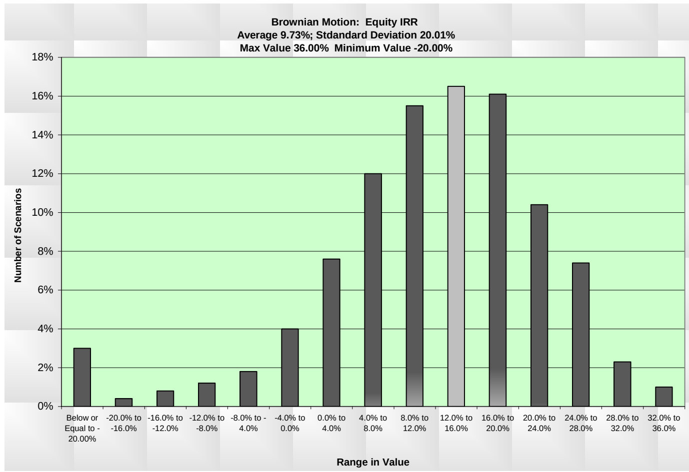
	Volatilty	Mean Reversion	Price Boundaries	Jump Process	Correlations	Probability of Total Equity Loss	Probability of Loss on Senior Debt	Probability of Loss on Subordinated Debt	Median Equity IRR	Median Project IRR	Probability of Negative Equity IRR	Probability of Negative Project IRR
Case 3	40%	None	None	None	None	32%	8.47%	24.89%	5.17%	5.05%	45%	37%
Case 7	40%	50%	None	None	None	2%	0.00%	0.00%	11.73%	8.83%	12%	3%
Case 8	40%	30%	None	None	None	10%	0.00%	0.40%	10.82%	7.83%	24%	14%
Case 9	40%	10%	None	None	None	25%	1.82%	7.60%	7.55%	6.34%	39%	30%
Case 10	40%	None	42	None	None	24%	0.72%	3.91%	7.54%	6.42%	38%	28%
Case 11	40%	10%	42	None	None	18%	0.16%	1.66%	9.03%	7.19%	33%	22%

Price Jumps

Instead of limiting price movements with price boundaries, one may want to model sudden price movements that can occur in economic series. The review of stock prices and oil prices above demonstrated that the percent changes did not follow a normal distribution in part because there were occasional large movements that would not be occur in a normal distribution. The case study presented in Chapter 1 demonstrated the dramatic price movements that occurred in California electricity markets. Sudden movements can occur because of an event such as a war, a bankruptcy filing a strike, a new technology innovation, a large new market entrant or a political decision. The parameters required to model a jump process include the probability of the jump occurring, the size of the jump, the standard deviation of the jump when it does occur, and the mean reversion factor for the jump. The jump process is driven by the “arrival rate” or the probability that the jump occurs in the equation. Variables representing the arrival rate and the size of the jump are driven by random draws as for other stochastic variables. If a jump process is included in the equation, assumptions are required to model what happens after the jump has occurred. If a jump process is modeled, then one of the key assumptions is the speed of mean reversion over which the jump is eliminated. The size of the jump can be expressed as the volatility multiplied by the base value multiplied by a number from 3 to 6 to represent the number of standard deviations from the mean.

For electricity prices, the jumps caused by capacity constraints last over relatively short periods driven by heat waves and/or plant outages. These jump processes clearly revert back to a mean level and do not remain in place permanently (as may be the case with a technological innovation or a new market entrant). Modeling future price jumps depends on how the probability of jumps and the size of the jumps is affected by new capacity additions and demand growth. Intuitively, it is simple to measure the effect of price jumps on the value of a plant. If a power plant has a 30 year life and there is a 10% probability of a price jump arriving at any year of the plant’s life, there are three expected jumps over the life of the plant.

Care must be taken in modeling price jumps in a time series analysis. If the market is competitive, the effect of the price jump on profits may be quickly eliminated by competitive pressure and new companies coming into the market. If prices that are faced by an investment increase because of price jumps, the effect is to increase returns rather than to increase the risks. Using the electricity case discussed throughout this section, a jump of four times the volatility is assumed to occur for .5% of the time. Even though the probability is low, the effect of the price jump is increase the possibility of earning high returns.



Correlations among Variables

Up to this point, the time series processes have been discussed for one variable without regard to how the movement is affected by changes in other variables. In most situations, analyzing a time series equation in isolation is inappropriate because key variables are correlated with another and a movement of one variable will affect the way other variables change. For example, in the case of the electricity plant fired with natural gas used to demonstrate the effect of alternative time series equations there would likely be a correlation between natural gas prices and electricity prices. If a time series equation is used in forecasting the price of electricity, the price of natural gas should have related movements in another time series model because of the ability to substitute fuels, because natural gas plants are often marginally running plants that drive the price and because the price of natural gas influence the type of new plants that are constructed. The relation between natural gas and electricity prices is so important that the difference between electricity price and natural gas price is often modeled and tracked separately from either the electricity price or the gas price. This difference in price is named the spark spread as shown on the formula below.

Spark Spread = Electricity Price ($/MWH) – Gas Price ($/MMBTU) x Heat Rate

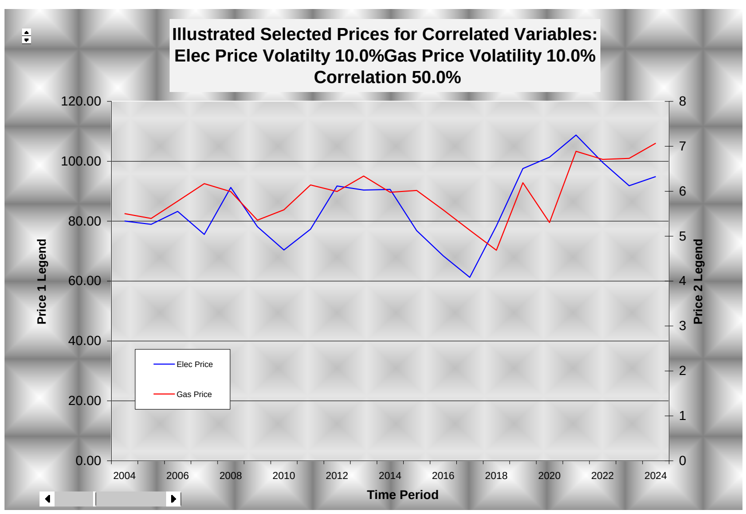
Using the spark spread to model the cash flow of a gas plant presumes that there is perfect correlation between gas and electricity prices. In some markets, there is very high correlation between natural gas and electricity prices but in other markets the correlation is relatively weak.

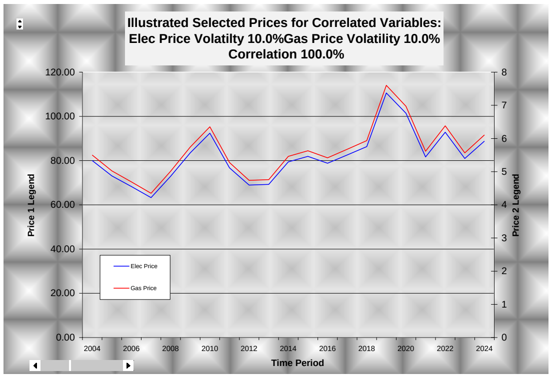
In the example above where there is lack of perfect correlation between natural gas prices and electricity prices, an alternative way to analyze distributions of cash flow for the gas plant is to create a time series models of both gas and electricity and to incorporate the correlation between the prices in the model. The process of incorporating the interrelationship between variables in time series models begins by defining a parameter that specifies the correlation between variables. The linear correlation coefficient has a value between 1 and -1 and is defined by the following formula:

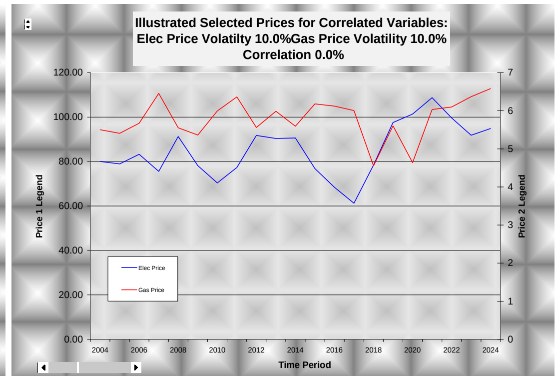
$$\text{Correlation}_{1,2} = [\text{Covariance}_{1,2}]/[(\text{Variance}_1 \times \text{Variance}_2)^{1/2}] = \rho_{1,2}$$

The correlation coefficient combined with volatility parameters can be used make a forecast of how the natural gas price and the electricity prices move together. The procedure for incorporating correlation into a time series model involves making random draw for the second correlated variable be a function of the random draw for the first variable. The mechanical procedure for making these calculations using something called the Cholesky function is somewhat dense and therefore relegated to Appendix B. This appendix attempts to explain in relatively simple terms the procedure for computing the random variation in one variable that is driven by the variation in another.

To illustrate the effect of reflecting correlation in variables in the analysis, the natural gas price in the earlier example is modeled as a stochastic variable rather than as a fixed amount as in the previous cases. Once reflecting volatility and mean reversion in the natural gas price, various different correlation factors are used to illustrate the effect of including correlation in the time series analysis. The effect of assuming no correlation, 50% correlation, 100% is evaluated in analysis illustrated below. For illustrating the effects of correlation, the case with 40% volatility and 10% mean reversion is used for both electricity and natural gas prices. In the first example, the natural gas prices are assumed to not be correlated with electricity prices, while in the second example, the correlation is assumed to be 50% and the in the third case the correlation is 100%. Selected price paths of the electricity prices and natural gas prices assuming different correlations (with volatility of 10%) are shown below. Note that in the case with 100% correlation the prices closely move together while in the case with 0% correlation there is no relation and in the case with 50% correlation, the prices generally move together, but the not all of the moves are coordinated.







The effects of correlation on financial statistics that measure risk illustrate the importance of the correlation assumption. In the case without correlation produces results similar to the case without any variability in gas prices. The case with no variability is analogous to debt leverage and aggravates risk. With positive correlation, the risks decline substantially as shown in the credit spread on subordinated debt column and the probability of a total loss on equity column. As with mean reversion, the table demonstrates that correlation is an important factor in measuring risk. Whether you ultimately attempt to make mathematical simulations or not, the exercises hopefully have demonstrated that thinking about whether an economic time series has mean reversion tendencies, boundaries, jumps or correlations with other variables is crucial in evaluating risk.

	Volatilty	Mean Reversion	Price Boundaries	Jump Process	Correlations	Probability of Total Equity Loss	Probability of Loss on Senior Debt	Probability of Loss on Subordinated Debt	Median Equity IRR	Median Project IRR	Probability of Negative Equity IRR	Probability of Negative Project IRR
Case 3	40%	None	None	None	None	32%	8.47%	24.89%	5.17%	5.05%	45%	37%
Case 7	40%	50%	None	None	None	2%	0.00%	0.00%	11.73%	8.83%	12%	3%
Case 8	40%	30%	None	None	None	10%	0.00%	0.40%	10.82%	7.83%	24%	14%
Case 9	40%	10%	None	None	None	25%	1.82%	7.60%	7.55%	6.34%	39%	30%
Case 12	40%	10%	0	None	0%	20%	1.30%	5.35%	11.27%	8.50%	32%	23%
Case 13	40%	10%	0	None	50%	5%	0.00%	0.10%	12.01%	9.01%	17%	7%
Case 14	40%	10%	0	None	100%	0%	0.00%	0.00%	11.76%	8.84%	2%	0%

Estimation of Time Series Parameters Using Statistical Analysis

You have probably noticed that none of the above discussion explained how to compute parameters for a time series equation. The exercises showed the importance of volatility, mean reversion, correlation and other variables, but not what the variables should be. Needless to say, if the construction of times series equations using parameters -- volatility, mean reversion, price boundaries, price jumps and correlations -- is to be useful in modeling, one must be able to calculate parameters for the models in an objective manner. This section discusses some of the practical and theoretical issues that arise when attempting to use statistical analysis of past data in computing various parameters. The discussion uses electricity and oil price data to demonstrate how the various calculations of volatility and mean reversion are made. Computing volatility parameters for a random walk process is described first. In subsequent sections, methods for computing volatility and mean reversion of a non-random walk process will be discussed.

The discussion focuses on parameters that should be input into a time series model rather than mechanical computation of statistics from historic data. To illustrate the difference, between input parameters and observed results, assume that a series of prices has a computed volatility of 20%, but that very tight upper and lower boundaries are modeled as part of the process. Because of the tight boundaries, the input parameter for volatility will not be the same as the resulting volatility in simulated cash flows.

Calculation of Volatility from a Random Walk Processes

Volatility is often defined as the standard deviation of annual percent changes in prices or some other variable. Since volatility is computed from the percent change rather than from absolute price levels, the unit of measurement for volatility is percentage – for example, volatility can be 20%, but it is would not be expressed as $30. Because volatility is expressed as an annual percentage rather than a daily or monthly percentage, if the standard deviation is computed from percentage changes in smaller time increments than annual increments, annualization adjustments are required. In the case of Brownian motion discussed below, for smaller increments than annual increments, the standard deviation is multiplied by the square root of the time increment.[9]

The three step procedure for calculating volatility of a random walk series without mean reversion involves first computing the rate of change in prices for the historic data. Second, the standard deviation of the series of rates of change is calculated. Finally, an adjustment is made for cases in which the rate of change is computed in time increments different from the time dimension of volatility (e.g. smaller than annual time increments.) The following formulas show how to compute volatility using both discrete and continual compounding with the three step process.[10] In both the discrete and continual cases, the first step is computing the rate of return over the period for reported prices. For continual compounding the rate of return is computed using the natural log:

$$\textit{Rate of Return}_i = \text{Natural Log}\ (\text{Price}_i/\text{Price}_{i-1})$$

Once a series of rates of return are established, standard deviation of the periodic rate of return is measured. If the prices are reported on an annual basis, the standard deviation of annual returns is volatility.

$$\textit{Period Volatility} = \text{Standard Deviation}\ (\textit{Rate of Return}_i)$$

The final part of the process for computing volatility where periodic prices are not expressed on an annual basis is converting the periodic volatility to an annual figure. Because of the mathematical process that defines Brownian motion, standard deviation of the rate of return increases with longer time periods. The variance of Brownian motion increases directly with time and the standard deviation increases with the square root of time. This means the period volatility defined above is multiplied by the square root of time measured in years (τ) to develop the annual volatility. Annual volatility is therefore defined as:

[9] For an explanation of why the standard deviation is multiplied by the square root of the time increment, see Pindyck, Robert and Dixit, Avinash, *Investment Under Uncertainty*, Princeton University Press, 1994, page 71.

[10] Section 2 of the workbook describes how to create a user defined function in spreadsheets that compute volatility whereby on can simply find the volatility of a series as one would find the average of a series.

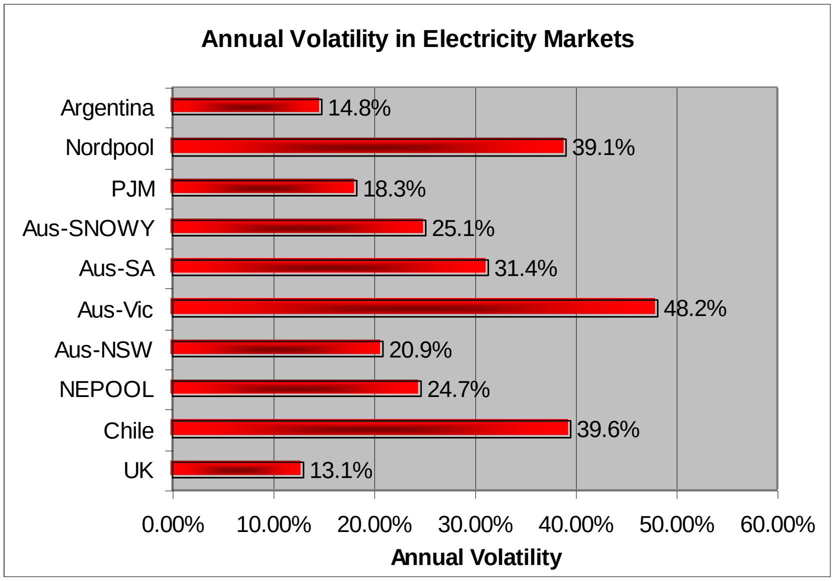
Annual Volatility = Standard Deviation (*Rate of Return*$_i$) x $(\tau)^{½}$

There are a variety of ways to compute the volatility of a time series that does not have mean reversion price boundaries or price jumps. One can compute volatility directly from historic price changes; one can compute the standard deviation of historic price changes divided by the average price; or one can use regression analysis of the change in price against the prior period price. Alternatively, if traded options exist for the variable in question, one can compute the volatility parameter that the market believes will occur in the future. The different measures of volatility for the S&P500 using different time periods are shown below:

In a random walk (Brownian motion) time series, the length of the time period used to measure volatility does not significantly impact the estimated volatility parameter. Computation of volatility is confirmed by inputting volatility into a Monte Carlo simulation and computing the standard deviation of the resulting distribution divided by the mean price and by the square root of the time increment. The standard deviation grows over time, but the standard deviation divided by the square root of time remains at the volatility level input.

A practical issue that often arises in analysis of time series is how much historical data should be used in establishing parameters of the equation. A general rule in statistics is that if the structure of the economic variables has not changed, as much data should be used as possible to maximize the number of degrees of freedom. If the structure of the market has changed, then data should be used since the change in structure. Of course, the problem is judging when a structural change resulting from new merchant capacity, changes in demand or different fuel prices has occurred of sufficient magnitude to warrant truncating the data set. The problem in resolving this data issue is one of the problems with relying solely on time series analysis in projecting future prices.

Volatility of annual electricity prices computed by measuring the standard deviation of annual rates of change is shown on the graph below. Volatility ranges from 13% for the UK to 48% for the State of Victoria in Australia. The volatility in the California market including the high prices in 2000 and 2001 was 101.6%. By contrast, the annual volatility in real oil and natural gas prices over the period 1976 to 2004 was 26.7% and 20.2% respectively.



The annual volatility must be consistent with the time series equations defined above meaning that over a large number of observations, volatility computed from the outcome of the model should be the same as the input. For example, if a variable does not have mean reversion, price spikes or price boundaries, and a volatility parameter of 20% is input into a model, the standard deviation of prices that are observed from historic data or modeled in a simulation should also be 20%. In the case of a random walk, the standard deviation of prices increases over time and the volatility parameter must reflect this fact. Volatility of a random walk time series can be estimated by computing the standard deviation of the percent change in price from one period to the next and adjusting the result for the length of the time period in historical data and the pricing model. Adjusting the standard deviation for the time increment is necessary because volatility is generally measured in annual terms and the standard deviation increases as more time passes in a random walk process.

Non-Constant High Volatility of Electricity

S&P 500 Volatility: 1994-2002				
Period		Standard Deviation of Rate of Return	Periods per Year	Annual Volatility Percent
Daily		1.19%	250	18.77%
Monthly		3.89%	12	13.47%
Annual		17.10%	1	17.10%

To demonstrate volatility properties of financial securities as compared to electricity prices, the S&P 500 index and New England electricity prices are evaluated. Volatility is computed in which returns are evaluated for prices in different time increments for example, daily prices, monthly average prices and annual average prices. Daily volatility is also computed for different periods such as different years, monthly periods within a year and semi-annual periods within a year. The tables below demonstrate

that when various different periods and/or time increments are used for the S&P 500, the volatility if quite stable. On the other hand when the same analysis is done for electricity prices, the volatility varies dramatically. Data for the volatility analysis is included in the reference files described in Section 5 of the workbook.

S&P 500 Volatility from Daily Returns			
Period	Standard Deviation of Daily Rate of Return	Periods per Year	Annual Volatility Percent
1997	1.14%	250	18.10%
1998	1.28%	250	20.30%
1999	1.14%	250	17.98%
2000	1.40%	250	22.14%
2001	1.36%	250	21.48%
2002	1.65%	250	26.09%
2001-First Half	1.45%	250	22.88%
2001-Second Half	1.21%	250	19.20%
2001-First Month	1.55%	250	24.49%
2001-Second Month	1.07%	250	16.91%
2001-Third Month	1.83%	250	29.00%
2001-Fourth Month	1.01%	250	15.98%
2001-Fifth Month	1.09%	250	17.31%
2001-Sixth Month	0.86%	250	13.61%

The first table demonstrates how standard deviation of the rate of return is converted to volatility for the S&P 500 and that even though the monthly volatility is lower than the daily volatility, the statistics are fairly stable. The second table demonstrates that if different time periods are used in evaluating volatility from daily returns, the statistics remain similar in the 20% range.

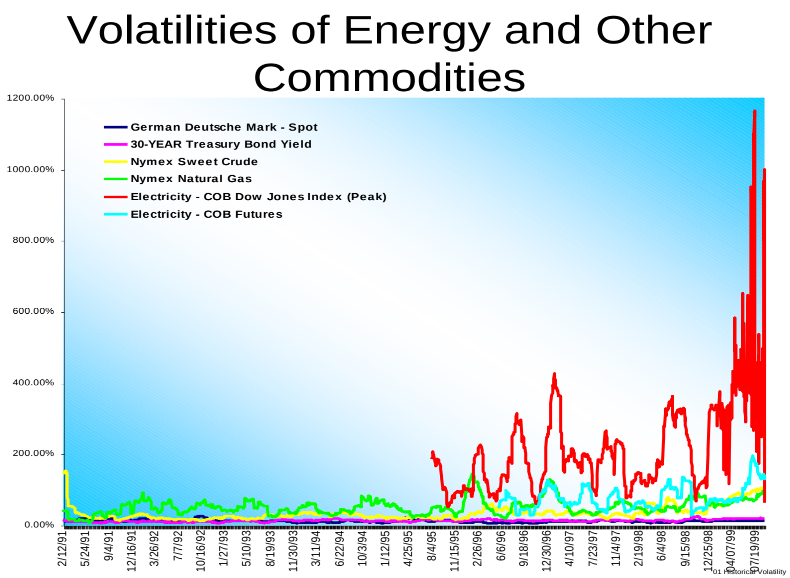
In a time series process with very high levels of volatility and mean reversion such as electricity, the volatility measured in short-term periods is much higher than in long term periods. The fact that volatility for electricity decreases with longer time frames of measurement has important implications on the structure of an option. If the strike provisions are based on monthly instead of hourly prices the volatility and therefore the value of an option are reduced significantly.

Using daily prices, annualized volatilities of electricity can approach 500%. Obviously, the annual change in electricity prices from one year to the next is not 500% (except in California.) The measures of short-term volatility for electricity are dramatically higher than estimated volatility for other commodities and financial securities. For example, the volatility of the S&P 500 index is about 10%, while the volatility of natural gas is about 50% and the volatility of crude oil is about 30%.[11] The insert compares the volatility of electricity to other components.

ENRON attempted to profit from trading in markets with high volatility. The following excerpt from ENRON's marketing materials illustrates this strategy: Electricity is by Far the Most Volatile of Commodities: Recent crude oil prices have seen price fluctuations as high as 70 percent. The volatility of coal spot prices over the past year in Europe has averaged 20 percent. Nickel, the most volatile metal traded, has seen volatility of 60 percent, and over the past five years, the price of natural gas futures has varied by 100 percent. However, none of these volatility figures compare to power prices this year in

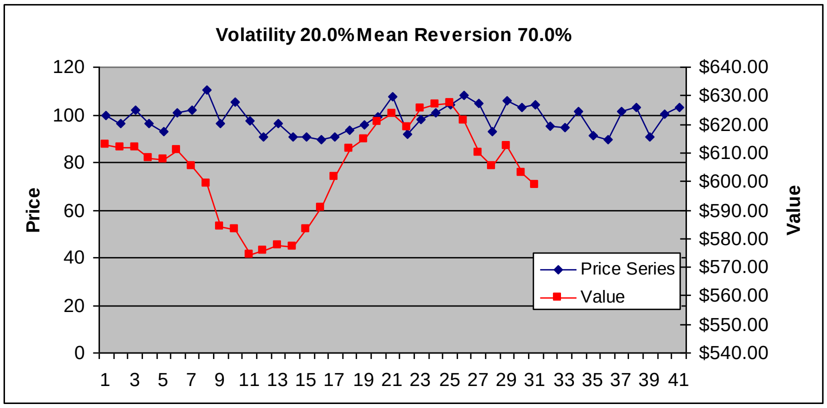
[11]Source: NYMEX presentation on futures contracts in 1995.

Europe. This January on the Amsterdam Power Exchange, the price of electricity suddenly jumped from 44 euros/MWh to 474 euros/MWh - a volatility of about 1,500 percent. The biggest factor at play in power markets is supply and demand. But because consumers don't know, in real time, how much electricity they are using, they can't react to high prices by cutting their consumption. According to Andersen Consulting, "Wholesale power markets, during periodic supply/demand crunches, have exhibited volatility 20 to 30 times that seen in financial and oil markets. This is driven by the fact that power cannot be truly stored."



Mean Reversion, Futures and Asset Valuation

While prices themselves follow a mean reverting process, this does not necessarily mean that securities which are derived from the present value of future mean reverting processes are also mean reverting. Consider the example of a share in a company the produces oil where the major uncertainty in cash flow comes from oil prices. Assume also that the oil price itself follows a mean reverting process. In this case, even though the oil price is mean reverting, movements in the share value follows a random walk process with no mean reversion. The random walk process occurs because the expectation of a mean reverting process is already priced in the current share value. Anytime there is a random shock in the oil price, the share value moves, but the percentage change in the share value is not as large as the percent change in the oil price. The relationship between price movement and value movement is demonstrated in the graph below (the source of this graph is the Monte Carlo exercise from section 2 of the workbook).



A similar phenomenon occurs in evaluating the characteristics of the prices of forward contracts and the price of the commodity underlying the forward contract. Even though the price of the commodity may follow a mean reverting process, the movement in the value of the forward contract over time has less mean reversion. If the forward contract only has term of one month, then the time series characteristics of the contract are similar to the price of the commodity itself. In this case, if there is mean reversion for the commodity, there is also mean reversion for the future contract. However, if the contract is for a longer term such as ten years, then moves in the commodity price are moderated because of the mean reversion in prices as was the case with the share value example above.

Discuss example with oil prices.

Method 1: Construction of historic volatility
- How long
- Where to get the data
- Adjustments for time period
- Implicit assumptions – that the rate of return follows a normal distribution

Method 2: Computation of implied volatility
- Find option
- Plug in values including strike price
- Use appropriate model
- Use goal seek with macro to find volatility

Relying on implied volatility alone is risky. Implied volatility simply tells you how options are currently priced, but not whether they are realistically priced. Historic volatility, on the other hand, can help you understand whether or not options are currently cheap or expensive.

If make simulation, see if the simulated volatility of each series is the same as the expected volatility. This falls apart when there is mean reversion, price boundaries and price jumps.

Computation of Time Series Parameters in Presence of Mean Reversion

Time period is crucial. Cannot apply monthly volatility to an annual model.

Computations of volatility are fairly straightforward if there is no mean reversion in the series. Where mean reversion is present, on the other hand, the calculations become more tricky. The first problem is that the mean reversion parameter measuring how long it takes for a series to come back to average levels is very difficult to extract from historic data. The second problem is that when volatility is computed from historic data and then used in a time series equation with mean reversion, the parameter that is input will not be the same as the resulting volatility output. The example below shows resulting volatility in a simulation where the input parameter for volatility is 20%. In the first case there is no mean reversion and in the second case mean reversion of 100% is assumed. The results are illustrated in the graphs below.

The general point of computing time series parameters is to derive statistics from historic data and then apply the statistics to mathematical models. If computation of a time series parameter from historic data is inconsistent with formulation of the mathematical model, this process will not work and the model will not produce unbiased results. For example, if 20% volatility is input into the model, on average, there should be 20% volatility in the simulated time series data that results from application of the model. With mean reversion and/or jumps and/or boundaries in models, the volatility formulas described above for a random walk series will not meet this proposition.

While the example produces the same results for both a mean reverting and a non-mean reverting series when the current price is the average price, the result does not hold when prices are different from the average. Consider a situation where the average price is 50, and the current price is still 100. In this case, a model driven by 20% volatility will produce different outcomes if the series does or does not include mean reversion. For the mean reversion series, the observed volatility from the 50 base price will be greater than 20% because the change in price includes the movement from 50 to 100 (a change of 100%) as well as effect of the 20% volatility. The time series without mean reversion will not have the added observed volatility driven by movements towards the mean and it will still have a 20% volatilty. Therefore, in the presence of mean reversion, if one measures volatility from the outcome of the prices using the standard deviation of price changes, the measured volatility will be more than 20% even though the volatility parameter driving the process is 20%. From the perspective of measuring volatility using observed historic prices, this means a computed volatility (say 100%) may be result from a time series process that has a lower true parameter driving the process (say 20%). In sum, if there is mean reversion, the volatility estimated from computing the standard deviation of the actual price changes will overstate the parameter required for the model.

In introducing mean reversion it was noted that computed volatility is not constant when measured over different time periods. Specifically, the volatility is greater when measured over short time periods than over long time periods. To demonstrate these problems consider a series with 100% mean reversion. Here, random movements away from the average are followed by movements back to the average level in the next period. Because prices keep coming back to the average level, the standard deviation of price changes does not increase over time as with a random walk series where prices wander indefinitely. This means that for mean reverting time series, one cannot easily estimate parameters using data with small time increments – days or months – and then derive annual parameters. If a time

series computed on a daily basis has high volatility (e.g. daily temperatures) the daily volatility cannot be used to derive annual volatility (e.g. the volatility of the annual average temperature is much lower than daily volatility).

The problems of measuring volatility arise even where time increments of the data are the same as time increments in the model – for example where annual data is used to estimate parameters and the model simulates annual movements in price. Say that the underlying process for a mean reverting series and a random walk series without mean reversion both are driven by 20% annual volatility and that the current price is 100. If the current price is also the average price, both cases would produce similar results for the subsequent year driven by the volatility parameter (there is a 68% chance that prices in the next period are between 80 and 120.) In this case, the observed volatility measured as the standard deviation of the percent change in price is the same as the volatility parameter input into the model for both the mean reverting series and the non-mean reverting time series. For both cases, after the 100 starting point, the next period moves up and down in a similar manner.

If one knew the mean reversion parameter beforehand, one could first compute the expected price with the effect of mean reversion. In the example above with mean reversion of 100%, volatility could be computed as the standard deviation between the average price and the realized price. Using the example above where the price is 50, the volatility is computed as the standard deviation between the average price of 100 rather than the last period price. The problem with this approach is that the mean reversion parameter is not known beforehand.

It has been suggested that mean reversion parameters can be computed from a regression equation. This approach however produces weak statistical results and is not consistent with simulations. Given that it is somewhat complex.

Construction of Time Series Models from Supply and Demand Models

If an estimate of long-run prices is required to assess the risks associated with an investment that may last from twenty to forty years, annual time series parameters constructed from historic data are likely to have little relevance. The volatility estimates for annual price data will be very different from the shorter term volatility parameters because, over the course of a year, weather reverts to normal levels, seasonality of maintenance schedules become less significant and other factors revert back to normal levels. Challenges arise in estimating annual volatility and particularly mean reversion for electricity prices because of the lack of historic price data for long periods and because there is less stability in the structure of the market (e.g. increases in capacity) the longer the time series. The next two chapters describe how long-run volatility can synthetically be computed using simulated data from supply and demand models. With supply and demand models, volatility is simulated from variation in loads, hydro conditions, fuel prices and maintenance outages.

Estimation of Other Parameters

In developing a time series model, a number of parameters other than volatility and mean reversion must often be estimated. Other parameters include time trends, equilibrium prices, lower and upper

boundaries, correlations among variables, price jumps and jump arrival rates as well as volatility of price boundaries, trends and equilibrium prices. With the exception of correlations, there is little objective historic data that can be used to estimate the parameters. Instead, judgmental considerations must be the basis for the parameters. For example, in determining lower boundaries, one must have knowledge of the production function of suppliers in the industry. Similarly, the long-term marginal cost requires knowledge of the cost of constructing different types of new facilities and the existing supply situation.

Parameters for simulating a price jump can be estimated from historic data if it is assumed that there will be no changes in the structure of the market (e.g. no capacity additions, demand growth, changes in maintenance schedules.) Modeling a jump process required four parameters – (1) the probability of a jump, (2) the average size of a jump, (3) the standard deviation of a jump, and (4) the mean reversion of the jump.The following step by step process demonstrates an approach for incorporating price jumps into a time series equation derived from historic data:

> Step 1: Establish a level above which a price spike is defined. For example, assume that any price above $50/MWH is classified as a price spike.
>
> Step 2: Use the assumed price spike levels to compute the number of price spikes and the average amount of the price spike in the historic data – the number of price jumps divided by the total number of periods is the probability of the jump occuring.
>
> Step 3: Compute the average level of the price jump and the standard deviation of price spikes. These are computed simply computing statistics from the actual price spike data.
>
> Step 4: Evaluate the mean reversion of price spikes by measuring the time period before which price spikes disappear. By the definition of a price spike, the mean reversion of price spikes should be very high.

Monte Carlo Simulation

Once parameters of a time series model have been estimated, a financial model must be developed so that risk can be measured for decision making. Financial models that use Monte Carlo simulation have the advantage that there is no need to do complex math and with Monte Carlo simulation one can visualize how prices are assumed to move over time with different model parameters.

The six step process above assumes that the distribution of prices follow a normal distribution. This assumption that price movements follow a normal distribution is not necessary. In converting the probabilies of making a particular random draw to a simulate price series, Monte Carlo simulation can be used with any type of probability distribution deemed appropriate to represent the nature of a price moves. Filtering random draws through a normal distribution implies that it is more likely for the variables to fall near the prior period price than far away from the prior period price. The graph below shows the standard normal distribution where the mean value is zero and the standard deviation is 1.0.

Appendix A:

Use of Continual Compounding Rather the Discrete Compounding in Time Series Equations

Finance and economic texts that describe Monte Carlo simulation and time series models often insist that it is proper construct simulations using continuious distributions rather than discrete distributions. If the assumption is made that returns come from a continuious distribution rather than assuming that changes occur at discrete intervals, some of the mathematical analysis becomes more confusing, but the basic ideas do not change. Those who insist on using continual distributions can be somewhat snobbish and scoff at models that assume discrete changes. The idea of this appendix is to work through differences between the continual and discrete equations so that you will not be intimidated by the differences.

To begin the discussion, consider the simple question of discrete versus continual compounding. The formulas for computing future values with different assumptions is shown in the table below. If the rate of return is relatively low (or the volatility is relatively low) the magnitude of differences is not very great.

Formulas for Continual and Discrete Compounding

Rate of Change	30%	
Beginning Value	100.000	
One Year - Discrete	130.000	=Beginning_Value*(1+Rate_of_Change)
One Year - Continuous	134.986	=Beginning_Value*EXP(Rate_of_Change)
Semi-Annual	132.250	=Beginning_Value*(1+Rate_of_Change/2)^2
Quarterly	133.547	=Beginning_Value*(1+Rate_of_Change/4)^4
Daily	134.969	=Beginning_Value*(1+Rate_of_Change/365)^365

If the rate of return is a large negative number, then the differences between continual and discrete compounding can be more dramatic. The table below shows that with a return of negative 300%, the discrete compounding case produces a negative number, while the continuous compounding case or the daily compounding case produces a small positive number.

Formulas for Continual and Discrete Compounding

Rate of Change	-300%	
Beginning Value	100.000	
One Year - Discrete	(200.000)	=Beginning_Value*(1+Rate_of_Change)
One Year - Continuous	4.979	=Beginning_Value*EXP(Rate_of_Change)
Semi-Annual	25.000	=Beginning_Value*(1+Rate_of_Change/2)^2
Quarterly	0.391	=Beginning_Value*(1+Rate_of_Change/4)^4
Daily	4.917	=Beginning_Value*(1+Rate_of_Change/365)^365

To understand why the continuous case produces a positive number, consider the case with daily compounding. In this case, the daily rate is -.82% meaning that on the first day of the year the 100 becomes 99.17. In the second period, this amount is reduces again and by the end of the year the base is reduced so that in the 364th day, the base is 4.958 resulting in the final number of 4.958 x (1-.82%) or 4.917. Because the returns are made smaller, the number continues to be positive. In time series analysis, the returns shown in the above table can be thought of as the volatility multiplied by the standard normal value. If the volatility is high – say 100%, and the normal draw is negative three, the outcome in the above example could occur.

The fact that prices cannot fall below zero with continual compounding is illustrated by the equation below where the percent change is computed from the natural log of the current price divided by the previous period price while the discrete change is computed from the standard formula (Current Price/Last Price – 1). In this equation, the term $e^{\text{Pct Change}}$ cannot be negative and therefore P_t also cannot be negative. The table below the formula demonstrates that when reductions occur, the measured percent change is above the discrete percent change – in absolute value terms. On the other hand, when prices are increasing, the measured continual rate of change is below the discrete measured change.

$$\text{Pct Change} = \text{Ln}(P_t / P_{t-1}), \text{ or } e^{\text{Pct Change}} = P_t / P_{t-1}, \text{ or } P_t = e^{\text{Pct Change}} \times P_{t-1}$$

Measured Percent with Discrete and Continuous Compounding

			Discrete	Continuous
Big Decline	Start	100		
	End	10	-90.00%	-230.26%
Increase			Discrete	Continuous
	Start	10		
	End	12	20.00%	18.23%
Decline			Discrete	Continuous
	Start	12		
	End	10	-16.67%	-18.23%

To see why the assumption of continual compounding makes time series equations a little more complicated, consider the case where the expected percent change in price is zero. If discrete equations are used, the likelihood of a lower price equals the likelihood of a higher price. For example, assume the volatility is 10%, meaning that the most likely percent change in price is zero, and there is a 68% change the price will increase or decrease by less than 10%. If the initial price is 100, and the percent change happens to be 10%, then the next price is 110. When the volatility is negative 10%, then the price is 90. The range in price on the upside and the downside is exactly the same. On the other hand, when continual compounding is used, then the expected price in the next period is not the same as the current price because the exponent of positive 10% does not result in the same price as the exponent of -10%. The table below illustrates the difference:

Base Price	100
Volatility	10%

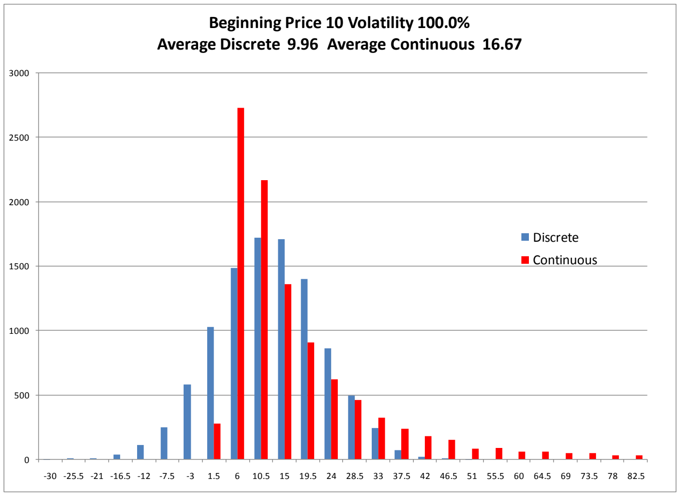
	Discrete			Continuous	
Sceanrio	Amount	Increase	Sceanrio	Amount	Increase
10%	110.00	10.00	10%	110.52	10.52
-10%	90.00	(10.00)	-10%	90.48	(9.52)
Average	100.00	0.00		100.50	0.50

While the small difference of .5 seems trivial, if the volatility is much higher, then the difference is much more as shown in the subsequent table which uses a volatility of 300% rather than 10%. Here the expected value difference is more than 900. To see why the difference is so large, return to the example of compounding with 365 days discussed above. Recall that the daily price change was .82%. When the price increased in the next period, then the .82% is applied to both the increase and the decrease. Similarly, if the price decreases, then the price increase or decrease is also applied. After going through 365 price increases or decreases, the low case is much less than the price increase case because the base of the price in the low case continues to decrease, while the base in the high case continues to increase.

Base Price	100
Volatility	300%

	Discrete			Continuous	
Sceanrio	Amount	Increase	Sceanrio	Amount	Increase
300%	400.00	300.00	300%	2,008.55	1,908.55
-300%	(200.00)	(300.00)	-300%	4.98	(95.02)
Average	100.00	-		1,006.77	906.77

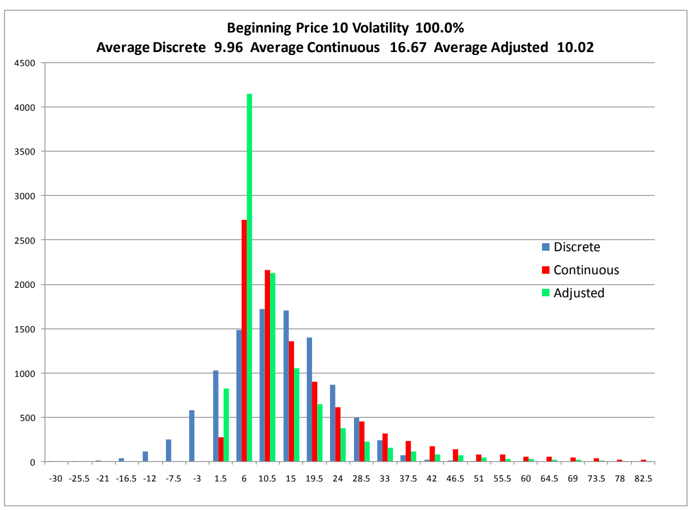
The above examples illustrated that the extent to which the average exceeds the initial price depends on the volatility. Further, the increase is non-linear. If 310% is used instead of 300%, the average increase is much more than .5 in the first table (it is more than 100.) A simulation can be easily constructed in excel to illustrate the effect of assuming a continual distribution versus a discrete distribution. Simply create a formula using the two alternative formulas – one with continual compounding using exp(volatility) and a second with discrete compounding using (1+volatility). Then multiply the volatility by a random draw from a normal distribution. To create the simulation, one can make a simple macro with a “for loop” and the “cells” function. A graph of price projections is shown below. The continual distribution cannot be negative and has a skewed distribution – a lognormal distribution -- while the discrete distribution has the same normal distribution as the percent change in prices.



In the above graph, the average price from the continual price assumption is 16.67 rather than 10, which is about the same as 10 x (e $^{(.5 \times (vol)^2)}$ -1). Specifically, the e $^{(.5 \times (vol)^2)}$ -1 is 6.49. This means that if we wanted to re-calibrate the expected price so that it still equaled the initial price, the term e $^{(.5 \times (vol)^2)}$ could be subtracted from each observation of the simulated price. If the general trend term is represented by α, then the time series equation with a time trend is:

$$P_t = P_{t-1} \times e^{(\alpha - .5 \times \text{volatility}^2 + \text{Standard Normal Draw} \times \text{Volatility})}.$$

The adjustment affects the distribution of prices as well as the average as shown on the graph below. Note that after the adjustment, the average of the prices is very close to the original base price of 10.



Appendix B:

Simulation of Correlated Variables

Many analyses require that the volatility and trends in more than one variable be incorporated in modeling analyses. For example, in evaluating the value at risk or the probability of default for projects sell or use include oil, natural gas and coal fuels, the risk of each of the variables must be analyzed. This part of the chapter describes how the Monte Carlo process can be implemented when there are multiple variables.

When simulating the potential distribution of more than one variable, if the variables are not correlated with each other, the time series equations can be modeled as independent equations where the standard normal draws are independent from one another. To create a model with a series of independent variables, Monte Carlo analyses described above can be separately performed for each variable. However, if the variables to be simulated are correlated with one another, running the single Monte Carlo over and over will not produce accurate results. Instead, one must model the variables simultaneously using a analytical approach based on something known as Cholesky factors.[12] Modeling oil and gas prices that are input into a supply and demand model requries the use of this process.

Cholesky factors are derived from matrix algebra to demonstrate that in the case where two variables are correlated with one another, the standard normal draw for the correlated variable should be the weighted combination of a new random draw and the random draw for the first variable. To illustrate how Cholesky factors are used recall that Monte Carlo simulation is driven by applying a volatility parameter to a factor -- z_t -- that is driven by a random variable. For the oil price, one first makes a transformation of random variables using the standard normal variable resulting in values approximately ranging from -4 to 4:

$$z_1 = \text{Inverse of Standard Normal}(u_1 \text{ for first variable})$$

To apply a similar factor to the second variable – the natural gas price. The second standard normal draw is defined as z_2:

$$z_2 = \text{Inverse of Standard Normal}(u_2)$$

After the second random draw is made and filtered through the normal distribution, an adjusted factor to apply to the volatility is computed from the correlation between the variables. The adjusted factor is a weighted average of the first two factors according to the following equation:

$$z\ adj_2 = z_1 \times \text{Correlation}_{1,2} + z_2 \times (1 - \text{Correlation}_{1,2}{}^2)^{1/2}$$

The weighting of the two random draws depends on the correlation between the variables. Two extreme cases demonstrate the way correlation is implemented in a Monte Carlo simulation framework. If there is no correlation between the variables, the factor for the second variable should not be influenced by the random draw for the first variable. Therefore, with no correlation the entire weight of the standard normal draw is given to the second random draw as expected. On the other hand, if the

[12] Jorion, Philippe, "Value at Risk", McGraw-Hill, 1997, pp. 242-243.

variables are perfectly correlated, the entire process should be driven by the first variable and there should be no weight given to the second random draw. This again is expected because any random shock that affects the first variable should also affect the second variable.

If three variables are correlated with each other instead of two variables, a similar process can be used. In this case, the correlation between the third variable and the second can be used in the equation and the equation does not require a memory of the random draw from the first variable. Since the second variable is already affected by the first variable, when the simulation process is finished, the first variable will be correlated with the third variable. Say the price of coal (variable 3) is correlated with the price of natural gas (variable 2) which in turn is correlated with the price of oil (variable 1). If the correlation between the price of gas and the price of coal is defined as $\rho_{2,3}$, the standard normal draw for the coal price is:

$$\text{Standard Normal Draw}_3 = \text{Standard Normal Draw}_2 \times \rho_{2,3} + \text{Inv Norm(New Random Draw)} \times (1 - \rho_{2,3}^2)^{1/2}$$

Here, the Standard Normal Draw_2 factor already incorporates the correlation between gas and oil prices.

Appendix C:

Use of Regression Analysis to Compute Mean Reversion and Volatility Parameters

An approach to computing volatility and mean reversion from historic data a regression can be used. The foundation for this approach is estimating the following equation[13]:

$$\text{Change in Price} = (P_t - P_{t-1}) = \alpha + \beta \times P_{t-1}$$

In this equation, the parameters α and β can be computed using a simple regression analysis of the change in price against the last period price. In this equation, there are some statistical problems with the β parameter which is biased towards zero (meaning that mean reversion is difficult to detect.) However, a reasonable estimate can be obtained and, if the parameter is significantly different from zero (i.e., the t statistic on β is above 2.0), the process is clearly mean reverting. In running the regression, there is no need to make fancy adjustments for autocorrelation or other complexities.

If the regression equation is estimated using annual data, estimated coefficients from the regression equation can be used to compute both the mean reversion parameter and the volatility parameter. If the data is expressed in monthly, daily or hourly terms, adjustments must be made to convert the regression

[13] Ibid, Dixtit and Pindic,

estimates to annual parameters. Before discussing how to use regression estimates of α, β and the standard error of the regression to establish the mean reversion and the volatility parameter, expected parameters of the regression equation are described.

Consider first the case where the β coefficient is equal to zero. In this case, the change in price has no relationship to the last period price. If the change in price is independent of historic prices, the equation meets a basic presumption of Brownian motion and a random walk. In a random walk, changes in price are not a function of past prices. When the β parameter is zero, the standard error of the regression is about the same as the standard deviation of the change in price – there is no variation in the constant term and the prior price drops out of the equation. This means that for an annual data series, the standard deviation of the change in price divided by the average price is the volatility. If the β parameter is significantly different from zero, the regression suggests that mean reversion is present in the time series. Here, in the case with mean reversion, the price change depends to a certain extent on the history of prices.

Estimation of Volatility and Mean Reversion from the Regression Equation

The objective of the statistical analysis discussed above is to ultimately develop a time series equation. This section describes how to convert coefficients of the regression equation into parameters of a time series model. The parameters include (1) the mean price (2) the mean reversion factor, and (3) the volatility.

Using the regression equation, the mean price is computed as:

$$\text{Mean Price} = -\alpha/\beta$$

The fact that this formula produces the mean price is demonstrated by simple algebra. Recall the formula for the change in price -- $\alpha + \beta \times P_{t-1}$. Over the sample period, the expected value of the change in price is zero. Further, the expected value of the last period price -- $E(P_{t-1})$ -- is the average price. Therefore, on an expected value basis, the formula reduces to:

$$0 = \alpha + \beta \times \text{Average Price},$$

that is rearranged to $-\alpha/\beta$ = Average Price.

The second parameter, the mean reversion factor, is computed from the β coefficient in the regression equation. Recall the mean reversion factor measures the average movement towards the mean price in a time period. This factor is defined as:

$$\text{Mean Reversion Factor} = -\log(1+\beta).$$

In this equation, if β is zero, the natural log of one minus zero is also zero, so the mean reversion parameter is zero. If β is -.63, the mean reversion factor is 1.0. If the β coefficient estimated from the regression equation is -.395, the mean reversion factor is .5 implying that in each period, prices move half way back to the mean level after a shock.

The third parameter of a time series equation that can be derived from the equation is the volatility. Volatility can be defined from the regression equation using the formula:

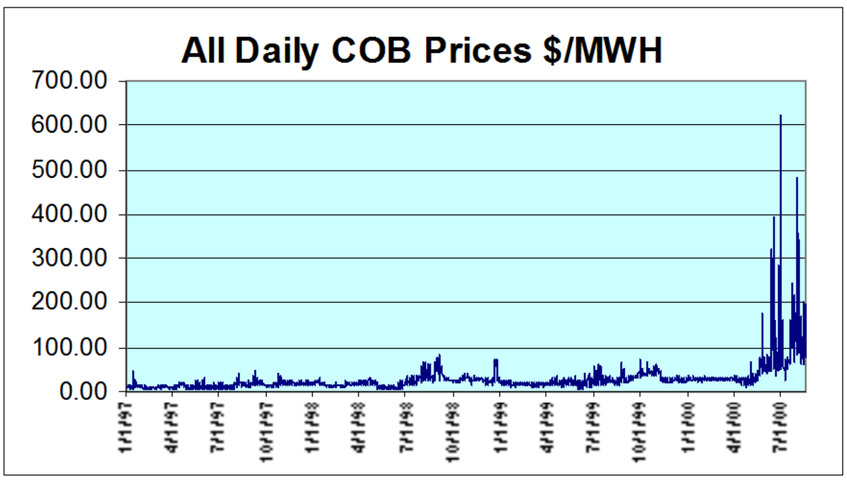
Absolute Volatility = Standard Error of Regression x $(\log(1+\beta)/((1+\beta)^2-1))^{(1/2)}$, and

Percent Volatility = Volatility/Average Price

In the above equation, if β is zero, the volatility level is the standard deviation of the change in price. The percent volatility is the standard error of the regression divided by the average price, or the standard deviation of the change in price divided by the average price. If β is greater than zero, the term $\log(1+\beta)$ is less than the term $(1+\beta)^2$, which means that the greater the β term, the smaller the volatility estimate. Intuitively, this means that if reversion to and back from the mean is causing some of the volatility, that volatility should be removed from the volatility that occurs exclusive of the mean reversion.

Calculating Mean Reversion Parameters

To illustrate calculation of parameters using the regression equation, this section reviews how a time series equation can be computed for prices at the California Oregon Border (COB) prior to the crisis in 2000. The graph below introduces the average daily on-peak and off-peak prices since 1997.



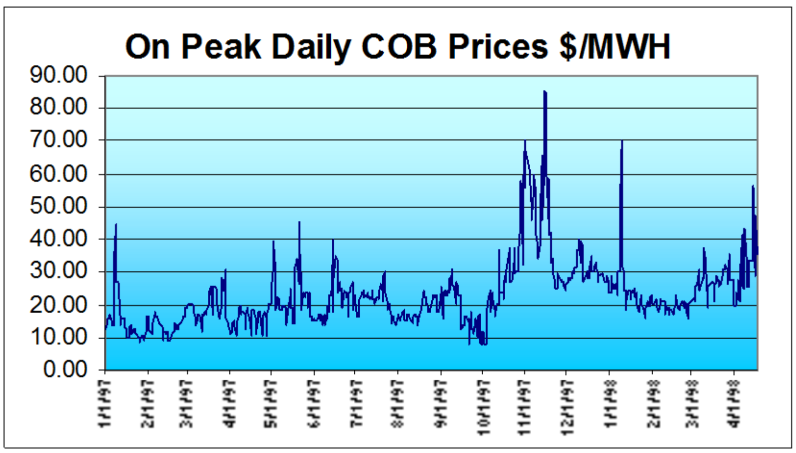
The accompanying graph demonstrates that price spikes began to occur in mid 2000, but before that, there were no prices above $100/MWH. To focus on computing parameters for the time series equations without the price spikes, only the data before 2000 is used in the statistical analysis.

While the formulas described above for the mean reversion factor, the mean prices and volatility seem straightforward, application to actual data raises a number of practical issues. Some of the issues involve when to separate models according to time period (summer versus non-summer and on-peak versus off-peak hours), how to account for price spikes and how much historical data to use in projections. Inspection of the data and knowledge of the markets clearly demonstrates that separate models should be developed for on-peak and off-peak prices. If one attempts to construct a model that includes both on-peak and off-peak prices, one will be trying to simulate a “saw tooth” pattern that will add nothing to the ability to make forecasts and will distort the statistics.

The graph above shows that in the summer of 2000, extended periods of very high prices occurred in both the on-peak and off-peak COB markets. These prices are a result of the much publicized power shortages in California that extended into 2001. Modeling the price levels after 2000 requires developing jump process parameters, which are discussed in a separate section. To focus the discussion on parameters that model price movements other than the price spikes, the remainder of this section concentrates on prices from 1997 through 1999. If there had been price spikes in the 1997-1999 data, we would have had to adjust the historic prices to subtract the price spikes.

Parameters for COB On-Peak Prices

The first series analyzed for purpose of developing time series parameters focuses on on-peak prices through the end of 1999. These prices are shown on the graph below. The graph demonstrates that there is mean reversion in the data and the lower bound is around $10/MWH. Instead of prices wandering aimlessly after a shock, the prices clearly move back to an average level. Every time the prices move up significantly, the prices later move back.



Using the COB on-peak price data, the regression analysis of the change in price versus the last period price produces estimates shown in the accompanying table. The t-statistic on the "last_price" variable is above 6 in absolute value, which suggests that the previous price has a strong relationship with the price change. This implies that from a statistical perspective, mean reversion exits.

```
UNWEIGHTED LEAST SQUARES LINEAR REGRESSION OF CHANGE

                                PREDICTOR
VARIABLES    COEFFICIENT    STD ERROR     STUDENT'S T       P
---------    -----------    ---------     -----------    ------
CONSTANT        2.31568       0.37047         6.25       0.0000
LAST_PRIC      -0.09108       0.01334        -6.83       0.0000

R-SQUARED           0.0458      RESID. MEAN SQUARE (MSE)     23.4575
ADJUSTED R-SQUARED  0.0448      STANDARD DEVIATION           4.84329

SOURCE        DF       SS           MS          F        P
----------   ---   ----------   ----------   -----   ------
REGRESSION     1     1093.99      1093.99    46.64   0.0000
RESIDUAL     972     22800.7      23.4575
TOTAL        973     23894.7
```

Parameters for Time Series Equations Using Alternative Time Periods

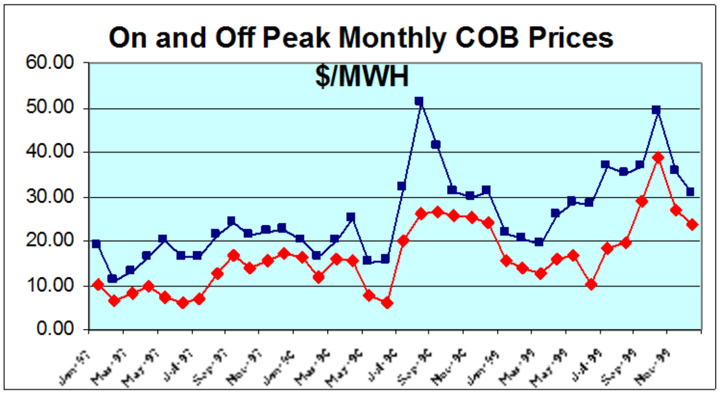
Estimated parameters differ depending on how the data is compiled. For example, monthly mean reversion differs from daily mean reversion and on peak mean reversion differs from off peak mean reversion. The volatility and the mean reversion parameters should be different for on-peak versus off-peak prices and for summer and non-summer prices. Parameters estimated from the regression equations for on-peak and off-peak time series are shown on the table below.[14]

	Off Peak Without Added Variables	Off Peak With Added Variables	On-Peak Summer	On-Peak Non-Summer
Regression Constant (a)	0.62404		3.28443	1.91129
Last Price Coefficient (b)	-0.03705	-0.07723	-0.11085	-0.08089
Standard Error of Regression	2.241	2.21711	6.60157	3.73169
Mean Price (-a/b)	16.84	16.84	16.84	16.84
Mean Reversion Factor (-log(1+b))	3.78%	8.04%	11.75%	8.43%
Multiplier (log(1+b)/((1+b)^2-1))^.5	0.720	0.736	0.749	0.737
Volatility	9.59%	9.68%	29.36%	16.33%
Annual Volatility	152%	153%	464%	258%

The table demonstrates that off-peak mean reversion and volatility is lower than for on-peak prices. Further, the volatility and mean reversion is less for the summer periods than for non-summer periods. This analysis suggests that separate models should be developed for alternative time periods. Section 2 of the workbook describes how to use spreadsheet techniques to efficiently compute volatility and mean reversion using spreadsheet techniques.

For some valuation issues, estimation of monthly price movements may be more relevant than projecting daily prices. For example, many purchase power contracts and tariffs have monthly prices rather than annual prices. As discussed above, for mean reverting time series, one cannot use data from daily data to project monthly parameters. Instead, monthly prices must be computed and then the parameters are computed from the monthly data. Monthly data for on-peak and off-peak COB prices are shown on the graph below.

[14] Annual volatility is presented using the square root of time as in the random walk method even though with mean reversion, the same time increment of the data should be used as the parameter for the time series model.



Regression analysis on the monthly prices demonstrates that the mean reversion and standard deviation are larger than with the daily prices. However, these larger numbers reflect a longer time period. Therefore, once prices are simulated, the deviation in prices may be more for the daily price series.

	Monthly On-Peak	Monthly Off-Peak
Regression Constant (a)	5.86202	2.8054
Last Price Coefficient (b)	-0.3429	-0.09676
Standard Error of Regression	6.55338	5.2408
Mean Price (-a/b)	17.10	28.99
Mean Reversion Factor (-log(1+b))	41.99%	10.18%
Multiplier (log(1+b)/((1+b)^2-1))^.5	0.860	0.743
Volatility	32.95%	13.44%
Annual Volatility	114%	47%

Problems with Time Series Equations for Analysis of Electricity Prices

This chapter introduced the characteristics of competitive electricity prices and it presented the mechanics of developing time series analysis. Before moving to other subjects, a few comments on problems with the times series analysis applicable to electricity are in order. Four issues and problems include:

(1) Time series analysis is more powerful when used together with supply and demand models than on a standalone basis in evaluating the risks associated with investment decisions;

(2) Regression analysis often does not produce adequate parameters for estimating time series equations when there is a high degree of mean reversion;

(3) Using statistical analysis of historic data in time series equations will produce erroneous analysis when structural changes have occurred in a market such as capacity additions or increased demand;

(4) When parameters are computed from price data without additional analysis, the underlying sources of volatility from price levels such as weather, hydro conditions, maintenance outages and economic activity can be ignored.

The statistical analysis used to develop time series parameters is not a simple matter of running a regression equation. Instead, the mean reversion parameter and the volatility must be tempered with judgment. Because of the limited ability to use objective statistical analysis in deriving parameters of the time series equations, analysts are often left trying alternative parameters and evaluating whether the outcome of price paths are reasonable.

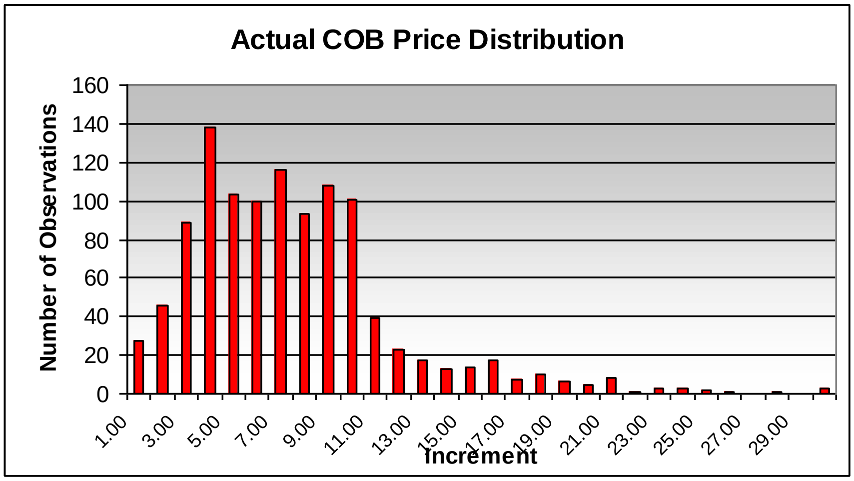
Adjusting parameters as described above is inadequate from an objective empirical standpoint. An alternative approach is to compare the distribution of simulated prices with the distribution of actual historic prices. The distribution of prices is computed by developing increments of price ranges and counting either simulated prices that fall within the range or by counting the number of actual occurrences that are in the range. An example of computing actual distributions is shown in the graph below. With the actual distribution computed one could test the simulation analysis with different parameters and compute the absolute difference between the simulated distribution and the absolute difference for each increment. This could be done on a relative basis as well as an actual basis as illustrated by the following formulas:

Actual Index = Actual Value/Average Value and Simulated Index = Simulated Value/Average Simulated

Using these index values, distributions are computed using the same bins for both the actual distribution and the simulated distribution as explained in section 2 of the workbook. Once these distributions are computed, the difference between actual and simulated distributions are computed and summed. The difference for each bin that is summed is:

$$\text{Difference}_{\text{Bin}} = \text{Absolute Value } (\text{Actual Index}_{\text{Bin}} - \text{Simulated Index}_{\text{Bin}})$$

Aside from the question of whether time series models adequately represent historic data, the models must be evaluated from a forecasting perspective. It is arguable that the massive financial problems that occurred from the demise of Long Term Capital Management arose from assumptions that the structure of markets – volatility and other parameters – could be estimated from historic data. If time series equations are entirely derived from historic data, the models cannot handle shifts in economic parameters resulting from changes in the structure of the market. For example, the addition of significant amounts of capacity in the market will reduce the probability of jumps and lower the volatility and mean price. Assessing how these parameters will change without delving into supply and demand models is virtually impossible.



A final problem with the time series analysis is that parameter estimation from actual prices ignores basic economic data on demand and supply drivers. For example, assume that it is known that prices were high because of abnormal weather or higher than normal plant outages. The analysis should not make prospective adjustments for normal conditions. Estimating parameters from actual prices does not do this.

Chapter Review

This chapter introduced price movements in electricity prices, technical definition of volatility, mean reversion and other parameters, Monte Carlo simulation, and statistical estimation of parameters. The following are the key points with respect to time series models of electricity prices:

(1) Market clearing prices of electricity follow a mean reverting process with punctuated jumps;

(2) Different time series models are appropriate for electricity during alternative seasons and time of day;

(3) Volatility parameters, rates of mean reversion, the boundary conditions and the price jumps can be derived from fundamental supply and demand models that simulate the underlying economic behavior of electricity prices;

(4) The specific parameters of time series models can be quantified with statistical analysis, but the statistics should be tempered with expert judgment about characteristics of markets.

Observation of actual market clearing prices of electricity demonstrates that prices are strongly mean reverting. Whether competitive electricity prices are measured on an hourly, daily or monthly basis, and whether the prices are measured in the New England region of the U.S., Scandinavia, or Australia, electricity generation prices follow certain common patterns. These patterns are characterized by sudden jumps (or less pronounced gradual movements) later followed by movement

back to the average price level. The nature of prices is different for different time periods and volatility is higher when prices are higher.

Parameters in time series models of electricity prices differ in alternative time periods. The basic demand and supply characteristics that drive prices in peak periods are different from the characteristics in off-peak periods. Similarly, the economic factors that drive prices in one season may not be the same as the factors that drive prices in another season. Instead of trying to develop models that incorporate seasonal and peak period price differences in a single equation, separate time series models can be constructed with different parameters. Further, the time series parameters and the underlying economic drivers are very different in developing short-term models that cover a few months or a year from a long-term model that is used to compute the value of a generating plant that may last for thirty to forty years.

Economic factors drive the rate of mean reversion in prices, boundary condition on prices and the price jumps in time series models. In constructing time series equations of electricity prices, those economic conditions that drive the price movements should be considered in an explicit or implicit manner. For example, the cost structure of electricity as represented by the supply curve, combined with historic volatility in demand can be used to simulate volatility parameters. Similarly, the behavior of independent suppliers in terms of profit on new plants drives reversion to mean "equilibrium" price levels in both the short-term and the long-run. The lower boundary condition on prices can be related to the short-run marginal cost of production and the jump processes can be derived from supply and demand models that explain prices during capacity constrained periods.

Parameters of time series models can be estimated through combining judgment with regression models. Basic inputs required for time series models include estimates of volatility and mean reversion. These parameters can be estimated with relatively simple statistical models. However, judgmental considerations can supplement the parameters and incorporate changes in the market structure. Judgment can also be used to include the probability and the severity of price jumps and the estimates of price boundary conditions. A problem with time series models that only use historic data and no judgment is that parameters in the models are difficult to test with statistical rigor. Instead, the analyst is relegated to "psychic" interpretations of trends in graphs.

Other industries that have deregulated, such as airlines, long-distance telephones, natural gas supply, banking and trucking have experienced significant reductions in price have occurred when price regulation was eliminated. However, the question of whether deregulation of electricity will ultimately produce significant price reductions for consumers is a murky picture -- some suggest that regulation contained massive inefficiencies and others suggest regulation artificially suppressed prices. While questions continue to exist regarding whether deregulation will ultimately increase or reduce prices, increases in the volatility of prices do generally accompany deregulation in industries that have moved from away from regulation.

Computer models and databases that accompany the analysis in this chapter are included in Section 2 of the workbook. The workbook includes databases of historic electricity prices from around the world; spreadsheet files that compute price volatility, mean reversion and correlation parameters from various different series; and, Monte Carlo simulation software that is used to make price forecasts. In addition, the accompanying Section 2 of the workbook describes spreadsheet techniques that can be used to manage hourly price data.

Time series analysis concepts introduced in this chapter will be used for various valuation and forward pricing issues addressed in subsequent chapters. For example, time series parameters such as volatility and mean reversion are used in forecasting electricity demands, hydro production, fuel prices and other supply and demand elements in Chapter 6 for predicting short-term price movements. Here, volatility parameters are applied to input variables to price models rather than prices themselves. Similarly, volatility is applied to the cost and productivity of new capacity as part of the modeling process for long-run price trends in Chapter 7. The mathematical concepts of volatility, mean reversion, price boundaries and price spikes are also the foundation for the option valuation models covered in Chapters 4, 6 and 7.